

## SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: \_\_\_\_\_ Examiner #: \_\_\_\_\_ Date: \_\_\_\_\_  
 Art Unit: \_\_\_\_\_ Phone Number 30 \_\_\_\_\_ Serial Number: \_\_\_\_\_  
 Mail Box and Bldg/Room Location: \_\_\_\_\_ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

\*\*\*\*\*

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc., if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: \_\_\_\_\_

Inventors (please provide full names): \_\_\_\_\_

Earliest Priority Filing Date: \_\_\_\_\_

*\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

## STAFF USE ONLY

Searcher \_\_\_\_\_  
 Searcher Phone #: **if Contact: Sheppard**  
 Searcher Location: \_\_\_\_\_  
 Date Searcher Picked Up: **tel: 308-4499**  
 Date Completed: **2/3/01**  
 Searcher Prep & Review Time: \_\_\_\_\_  
 Clerical Prep Time: \_\_\_\_\_  
 Online Time: \_\_\_\_\_

## Type of Search

NA Sequence (#) \_\_\_\_\_ STN \_\_\_\_\_  
 AA Sequence (#) \_\_\_\_\_ Dialog \_\_\_\_\_  
 Structure (#) \_\_\_\_\_ Questel/Orbit \_\_\_\_\_  
 Bibliographic \_\_\_\_\_ Dr. Link \_\_\_\_\_  
 Litigation \_\_\_\_\_ Lexis/Nexis \_\_\_\_\_  
 Fulltext \_\_\_\_\_ Sequence Systems \_\_\_\_\_  
 Patent Family \_\_\_\_\_ WWW/Internet \_\_\_\_\_  
 Other \_\_\_\_\_ Other (specify) \_\_\_\_\_

## Vendors and cost where applicable

---

**THIS PAGE BLANK (USPTO)**

Result	No.	Score	Query	Match	Length	DB	ID	Description
1	296	81.3	226	12	009727		009727	leonorus mo
2	296	81.3	226	12	09WHf6		09WHf6	tomato mild
3	294	80.8	361	12	067574		067574	bean golden
4	292	80.2	225	12	Q9Q0B1		Q9Q0B1	cowpea gold
5	292	80.2	354	12	Q9Q5S5		Q9Q5S5	sweet potato
6	281	77.3	119	12	Q9H972		Q9H972	aida golden
7	281	77.3	119	12	088972		088972	aida golden
8	281	77.2	233	12	09YL44		09YL44	macroptiliu
9	275	75.5	234	12	039180		039180	potato yell
10	272	74.7	190	12	032089		032089	tobacco tea
11	272	74.7	190	12	032084		032084	tobacco tea
12	271	74.5	150	12	09WH27		09WH27	tobacco tea
13	271	74.5	208	12	0320C4		0320C4	tobacco tea
14	269	73.9	208	12	0320C0		0320C0	tobacco tea
15	269	73.9	208	12	Q320B8		Q320B8	tobacco tea
16	269	73.9	208	12	Q320B6		Q320B6	tobacco tea
17	266	73.0	203	12	0320B6		0320B6	tobacco leaf
18	266	73.1	363	12	073577		073577	cotton leaf
19	259	71.2	208	12	0320A0		0320A0	tobacco leaf

RN	{11}
RP	SEQUENCE FROM N.A.

RN	{11}
RP	SEQUENCE FROM N.A.



RESULT	7	
P88975		
ID	P88975	PRELIMINARY; PRT; 149 AA.
AC	P88975	
DT	01-MAY-1997	(T=EMBLrel. 03, Created)
DT	01-MAY-1997	(T=EMBLrel. 03, Last sequence update)

GN	REP.	Macroptilium golden mosaic virus.
NS	AC	Viruses: ssRNA viruses: geminiviridae: Begomovirus.
OC	OC	11: TAA10-15676; .
OC	RM	11: TAA10-15676; .
OC	RM	SEQUENCE FROM N.A.
RP	RP	STRAIN=JAMAICA STRAIN 1;
RP	RP	ROYE M.E.;
RT	RT	"Genetic diversity and phylogeny of whitefly-transmitted geminiviruses
RT	RT	from Jamaica."; 1998) to the EMBL/GenBank/DBJ databases.
RT	RT	121
RT	RT	SEQUENCE FROM N.A.
RP	RP	STRAIN=JAMAICA STRAIN 1;
RP	RP	ROYE M.E., McLaughlin W.A., Maxwell D.P.;
RT	RT	"Molecular characterization of two distinct geminiviruses infecting M.
RT	RT	Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.
RL	RL	EMBL: AF09840; XAD17850.1; .
DR	DR	INTERPRO: IPR001191; .
DR	DR	INTERPRO: IPR001301; .
DR	DR	PFAM: PF00759; GeminA_ALL; 1.
DR	DR	PRINTS: PR00228; GEMINATLVL.
DR	DR	PRINTS: PR00228; GEMINATLVL.
DR	DR	NON-TELT 233
DR	DR	SEQUENCE 233 AA; 26355 MW; AA490AFD2166A02 CRC64;
FT	FT	
SQ	SQ	

Query Match	77.2%;	Score 281;	DB 12;	Length 233;
Best Local Similarity	72.9%;	Pred. No. 4.5e-23;		

Matches 51; Conservative 12; Mismatches 7; Indels 0; Gaps 0;

QY 1 TLVWKEFQVGRSARGGQTSNDAAALNASSKEALQIREKIPEKYLEFQFINLSNL 60  
 DB 110 TIEMGVQIDGRSARGGQTSNDAAALNASSKEALQIREKIPEKYLEFQFINLSNL 169

QY 61 DRIFDKTPPEP 70  
 DB 170 DRIFDKTPPEP 179

RESULT 9  
 ID Q93180 PRELIMINARY: PRT: 234 AA.  
 AC Q93180;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 13, Last annotation update)  
 DE C1 N-TERMINUS: CURE YGSHAS-2 (FRAGMENT).  
 OS potato yellow mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 RN NCBI\_TaxID=10827;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-TOMATO STRAIN;  
 RA Guzman P., Arredondo C.R., Emattty D., Portillo R.J., Gilbertson R.L.;  
 RE EMBL: A5026553; AAB82605.1; -  
 DR EMBL: A5026553; AAB82605.1; -  
 DR INTERPRO: IPR001191; -  
 DR INTERPRO: IPR001301; -  
 DR PFAM: PF00799; Gemin\_ALI; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR PRINTS: PR00228; GEMCOATCLVL.  
 FT NON\_TER 234  
 FT SEQUENCE 234 AA; 26486 MW; 9ED8F0697105CD19 CRC64;

Query Match 75.5%; Score 275; DB 12; Length 234;  
 Best Local Similarity 71.4%; Pred. No. 2e-22;  
 Matches 50; Conservative 13; Mismatches 7; Indels 0; Gaps 0;

QY 1 TLVWKEFQVGRSARGGQTSNDAAALNASSKEALQIREKIPEKYLEFQFINLSNL 60  
 DB 110 TIEMGVQIDGRSARGGQTSNDAAALNASSKEALQIREKIPEKYLEFQFINLSNL 169

QY 61 DRIFDKTPPEP 70  
 DB 170 DRIFDKTPPEP 179

RESULT 10  
 ID Q92089 PRELIMINARY: PRT: 190 AA.  
 AC Q92089;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE C1 N-TERMINUS: CURE YGSHAS-2 (FRAGMENT).  
 OS tobacco leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 RN NCBI\_TaxID=67762;  
 RX [1]  
 RC STRAIN-FROM N.A.  
 RA Ooi K., Ohshita S., Ishii I., Yahara T.;  
 RE EMBL: A8001315; BAA34033.1; -  
 DR EMBL: A8001315; BAA34033.1; -  
 DR INTERPRO: IPR001191; -  
 DR INTERPRO: IPR001301; -  
 DR PFAM: PF00799; Gemin\_ALI; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR PRINTS: PR00228; GEMCOATCLVL.  
 FT NON\_TER 190  
 FT SEQUENCE 190 AA; 21444 MW; AACIC2943E3F01AD CRC64;

Query Match 74.7%; Score 272; DB 12; Length 190;  
 Best Local Similarity 62.4%; Pred. No. 3.4e-22;  
 Matches 53; Conservative 8; Mismatches 8; Indels 16; Gaps 1;

QY 1 TLVWKEFQVGRSARGGQTSNDAAALNASSKEALQIREKIPEKYLEFQFINLSNL 60  
 DB 85 TLEWGTQIDGRSARGGQTSNDAAALNASSKEALQIREKIPEKYLEFQFINLSNL 144

QY 61 DRIFDKTPPEP 70  
 DB 145 DRIFDKTPPEP 144

RESULT 12  
 ID Q9827 PRELIMINARY: PRT: 190 AA.  
 AC Q9827;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DE C1 N-TERMINUS: CURE YGSHAS-2 (FRAGMENT).  
 OS tobacco leaf curl virus.

DR PRINTS: PR00227; GEMCOATALL.  
 DR PRINTS: PR00228; GEMCOATCLVL.  
 FT NON\_TER 190  
 FT SEQUENCE 190 AA; 21432 MW; AAC093D1D1610PAD CRC64;

Query Match 74.7%; Score 272; DB 12; Length 190;  
 Best Local Similarity 62.4%; Pred. No. 3.4e-22;  
 Matches 53; Conservative 8; Mismatches 8; Indels 16; Gaps 1;

QY 1 TLVWKEFQVGRSARGGQTSNDAAALNASSKEALQIREKIPEKYLEFQFINLSNL 60  
 DB 85 TLEWGTQIDGRSARGGQTSNDAAALNASSKEALQIREKIPEKYLEFQFINLSNL 144

QY 61 DRIFDKTPPEP 70  
 DB 145 DRIFDKTPPEP 144

RESULT 11  
 ID Q92084 PRELIMINARY: PRT: 190 AA.  
 AC Q92084;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE C1 N-TERMINUS: CURE YGSHAS-2 (FRAGMENT).  
 OS tobacco leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 RN NCBI\_TaxID=67762;  
 RX [1]  
 RC STRAIN-FROM N.A.  
 RA Ooi K., Ohshita S., Ishii I., Yahara T.;  
 RE EMBL: A8001318; BAA34039.1; -  
 DR EMBL: A8001318; BAA34039.1; -  
 DR INTERPRO: IPR001191; -  
 DR INTERPRO: IPR001301; -  
 DR PFAM: PF00799; Gemin\_ALI; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR PRINTS: PR00228; GEMCOATCLVL.  
 FT NON\_TER 190  
 FT SEQUENCE 190 AA; 21444 MW; AACIC2943E3F01AD CRC64;

Query Match 74.7%; Score 272; DB 12; Length 190;  
 Best Local Similarity 62.4%; Pred. No. 3.4e-22;  
 Matches 53; Conservative 8; Mismatches 8; Indels 16; Gaps 1;

QY 1 TLVWKEFQVGRSARGGQTSNDAAALNASSKEALQIREKIPEKYLEFQFINLSNL 60  
 DB 85 TLEWGTQIDGRSARGGQTSNDAAALNASSKEALQIREKIPEKYLEFQFINLSNL 144

QY 61 DRIFDKTPPEP 70  
 DB 145 DRIFDKTPPEP 144

RESULT 12  
 ID Q9827 PRELIMINARY: PRT: 190 AA.  
 AC Q9827;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DE C1 N-TERMINUS: CURE YGSHAS-2 (FRAGMENT).  
 OS tobacco leaf curl virus.

OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.

```

OK NCBI_TaxID=67762;
RN SC "ABU33994.1"
RC STRAIN=ABU33994.1
PA Ooi K., Ohshita S., Ishii I., Yahara T.;
RT "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL J. Plant Res. 110:247-257(1997).
DR EMBL: AB001303; BAA34010.1; -.
DR INTERPRO: IPR001191; -.
DR PFAM: PF00799; Geminini_1.
DR PRINTS: PR00228; GEMCOATCLVL.
FT NON_TER 1 190
FT NON_TER 190 1
FT NON_TER 190 190
SQ SEQUENCE 190 AA; 21444 MW; 93C3742A8EB0B7EB CRC64;

```

```

Query Match      74.5%; Score 271; DB 12; Length 190;
Best Local Similarity 62.4%; Pred.No. 4.9e-22;
Matches 53; Conservative 8; Mismatches 16; Gaps 1;

Oy 1 TLVWGFEVQVGSARGCGCTGNDAAALNASSKEEALQIIRKIPKYLQFPHLNSNL 60
Db 85 TLVWGFEVQVGSARGCGCTGNDAAALNASSKEEALQIIRKIPKYLQFPHLNSNL 144
Oy 61 DRI-----FQKTPK 69
Db 145 DRIAPPLEVFVCPFSSSSFQVPE 169

```

```

RESULT 13
OQ20C4
ID Q320C4 PRELIMINARY; PRT; 208 AA.
AC Q320C4
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE C1 AND C4 GENES, CLONE ABU33-1, PARTIAL
DE AND COMPLETE CDS (FRAGMENT).
GN Cl.
OS tobacco leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=67762;
RN SC "ABU33994.1"
RC STRAIN=ABU33994.1
PA Ooi K., Ohshita S., Ishii I., Yahara T.;
RT "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL J. Plant Res. 110:247-257(1997).
DR EMBL: AB001303; BAA33994.1; -.
DR INTERPRO: IPR001191; -.
DR PFAM: PF00799; Geminini_1.
DR PRINTS: PR00227; GEMCOATCLVL.
DR PRINTS: PR00228; GEMCOATCLVL.
FT NON_TER 1 208
FT NON_TER 208 1
FT NON_TER 208 208
SQ SEQUENCE 208 AA; 23526 MW; 249CC31D8729C72D CRC64;

```

```

Query Match      74.5%; Score 271; DB 12; Length 208;
Best Local Similarity 62.4%; Pred.No. 4.9e-22;
Matches 53; Conservative 8; Mismatches 16; Gaps 1;

Oy 1 TLVWGFEVQVGSARGCGCTGNDAAALNASSKEEALQIIRKIPKYLQFPHLNSNL 60
Db 102 TLVWGFEVQVGSARGCGCTGNDAAALNASSKEEALQIIRKIPKYLQFPHLNSNL 161
Oy 61 DRI-----FQKTPK 69
Db 162 DRIAPPLEVFVCPFSSSSFQVPE 186

```

```

RESULT 14
OQ20C0
ID Q320C0 PRELIMINARY; PRT; 208 AA.
AC Q320C0
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE C1 AND C4 GENES, CLONE AMG-1(8152),
DE PARTIAL AND COMPLETE CDS (FRAGMENT).
GN Cl.
OS tobacco leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=67762;
RN SC "ABU33994.1"
RC STRAIN=ABU33994.1
PA Ooi K., Ohshita S., Ishii I., Yahara T.;
RT "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL J. Plant Res. 110:247-257(1997).
DR EMBL: AB001256; BAA33996.1; -.
DR INTERPRO: IPR001191; -.
DR PFAM: PF00799; Geminini_1.
DR PRINTS: PR00227; GEMCOATCLVL.
DR PRINTS: PR00228; GEMCOATCLVL.
FT NON_TER 1 208
FT NON_TER 208 1
FT NON_TER 208 208
SQ SEQUENCE 208 AA; 23486 MW; E301135F799C3DAD CRC64;

```

```

Query Match      73.9%; Score 269; DB 12; Length 208;
Best Local Similarity 61.2%; Pred.No. 8.1e-22;
Matches 52; Conservative 9; Mismatches 8; Indels 16; Gaps 1;

Oy 1 TLVWGFEVQVGSARGCGCTGNDAAALNASSKEEALQIIRKIPKYLQFPHLNSNL 60
Db 102 TLVWGFEVQVGSARGCGCTGNDAAALNASSKEEALQIIRKIPKYLQFPHLNSNL 161
Oy 61 DRI-----FQKTPK 69
Db 162 DRIAPPLEVFVCPFSSSSFQVPE 186

```

```

RESULT 15
OQ20B8
ID Q320B8 PRELIMINARY; PRT; 208 AA.
AC Q320B8
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE C1 AND C4 GENES, CLONE AMG-2(8154),
DE PARTIAL AND COMPLETE CDS (FRAGMENT).
GN Cl.
OS tobacco leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=67762;
RN SC "ABU33994.1"
RC STRAIN=ABU33994.1
PA Ooi K., Ohshita S., Ishii I., Yahara T.;
RT "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL J. Plant Res. 110:247-257(1997).
DR EMBL: AB001257; BAA33998.1; -.
DR INTERPRO: IPR001191; -.
DR PFAM: PF00799; Geminini_1.
DR PRINTS: PR00227; GEMCOATCLVL.
DR PRINTS: PR00228; GEMCOATCLVL.
FT NON_TER 1 208
FT NON_TER 208 1
FT NON_TER 208 208
SQ SEQUENCE 208 AA; 23472 MW; 629D0DEFC9956AA CRC64;

```

```

Query Match      73.9%; Score 269; DB 12; Length 208;
Local Similarity 61.2%; Ident No. 8; Id-22;
Matches 52; Conservative 9; Mismatches 6; Indels 16; Gaps 1;

Oy 1 TLVNGEFOVGRSARGGCGTSDNDAAEALNASSKEEALQIIREKIPKYLFOFHNLNSNL 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 102 TLEWGTFOIDGRSARGGCGQANDACAEALNASSKADALAIIREKLPKDFIQYHNLNSNL 161
Oy 61 DRI-----FDKTPP 69
   |||||
Db 162 DRIAPPLEVFVCFPSSSSDQVPE 186

```

Search completed: February 3, 2001, 02:22:50  
Job time: 1887 sec

Result	No.	Score	Query	Match	Length	DB	ID	Description
1	364	100.0	352	1	VALL_TGMY	P03567	tomato gold	
2	264	72.5	361	1	VALL_PVYV	P07258	potato yell	
3	251	69.0	358	1	VALL_CDKV	P14982	cassava yell	
4	251	67.0	365	1	VALL_CDKV	P03672	cassava lat	
5	221	67.0	359	1	VALL_TYLCU	P03629	tomato yell	
6	237	65.1	369	1	VALL_TYLCU	P38609	tomato yell	
7	233	64.0	349	1	VALL_PVUV	P06923	pepper huas	
8	230	63.2	359	1	VALL_TYLCU	P07260	tomato yell	
9	228	62.6	353	1	VALL_BGMV	P05175	bean golden	
10	227	62.4	355	1	VALL_ABMY	P21947	abutilon mo	
11	224	61.5	358	1	VALL_PCRV	P14991	beet curly	
12	224	61.3	361	1	VALL_PCRV	P06557	tomato mott	
13	224	61.3	361	1	VALL_PCRV	P06557	tomato yell	
14	124	34.1	347	1	VALL_SILV	P29048	squash	
15	64.5	17.7	299	1	Y175_HELJ	Q920q7	hellicobact	
16	64.5	17.7	1713	1	LMJ3_HESU	Q16787	homo sapien	
17	63.5	17.4	1610	1	CCAD_MEAN	ME003126	mesocricetu	
18	63.5	17.4	2161	1	CCAD_HUMAN	Q01668	homo sapien	
19	63.5	17.4	2203	1	CCAD_FERT	P277372	rattus norv	
20	62	17.0	435	1	VALL_TYDVA	P1617	tobacco yell	
21	60	16.5	447	1	CUSA_DROME	AA024095	drosophila	
22	60	16.5	447	1	CUSA_DROME	P56079	rhizobium s	
23	59	16.2	387	1	Y4PE_RHLSN	P55615	rhizobium s	
24	58.5	16.1	136	1	K452_CASL	Q62250	caenorhabdi	
25	58.5	16.1	297	1	RRP3_RABVP	P06747	rabies viru	
26	58.5	16.1	299	1	Y175_HELJ	P56112	hellicobact	
27	58.5	16.1	428	1	GBA1_CANAL	P28868	candida alb	
28	58.5	16.1	589	1	Y667_HACIN	P54221	hemophilus	
29	58	15.9	677	1	YACI_ECOLI	P36682	escherichia	
30	58	15.9	677	1	YACI_ECOLI	P36682	escherichia	
31	57.5	15.8	534	1	G3P_BACCO	P03652	bacillus th	
32	57.5	15.8	335	1	G3P_BACCO	P51115	bacillus co	
33	57.5	15.8	367	1	LHX1_MOUSE	P53776	mus musculu	

P27258;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DE ALI PROTEIN (Rel. 23, Last annotation update)  
 OS Potato yellow mosaic virus (isolate Venezuela).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 RN ([1])  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-91311403; PubMed-1856690.  
 RA Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.:  
 RT The nucleotide sequence of the infectious cloned DNA components of  
 RT potato yellow mosaic virus.  
 RL J. Gen. Virol. 72:1515-1520(1991).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))  
 CC  
 CC EMBL: D00940; BAA00782.1; -  
 DR PIR: JUD364; OCVPT.  
 DR INTERPRO: IPR001191;  
 DR PFAM: PF00749; Gemin\_A1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATCLV1.  
 KW ATP-binding.  
 FT NP\_BIND 222 229 ATP (POTENTIAL).  
 SQ SEQUENCE 361 AA: 40850 MW; 5627A33BF1264383 CRC64;

Query Match 72.5%; Score 264; DB 1; Length 361;  
 Best Local Similarity 71.0%; Pred. No. 2e-22;  
 Matches 49; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

QY 1 TLWGEFQVDSRGCGQTSNDAAEALNASSKEALQIREKIPKYLFOFHNLNSL 60  
 DB 110 TVMGQFQIDGSRGCGQTSNDAAEALNASSKEALQIREKIPKYLFOFHNLNSL 169  
 QY 61 DRIFDKTPE 69  
 DB 170 DRIFDKAPE 178

RESULT 3  
 ID VALL CLVW STANDARD; PRT: 358 AA.  
 AC P14982;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DE 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE ALI PROTEIN (40.4 KDA PROTEIN).  
 OS Cassava latent virus (strain West Konyak 844).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 RN ([1])  
 RP SEQUENCE FROM N.A.  
 RA Stanley J., Gay M.;  
 RT Nucleotide sequence of cassava latent virus DNA.\*;  
 RL Nature 301:260-262(1983).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: J02057; -; NOT\_ANNOTATED\_CDS.  
 DR INTERPRO: IPR001191;  
 DR PFAM: PF00749; Gemin\_A1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATCLV1.  
 KW ATP-binding.  
 FT NP\_BIND 220 227 ATP (POTENTIAL).  
 SQ SEQUENCE 358 AA: 40345 MW; ED173E753EE92069 CRC64;

Query Match 69.0%; Score 251; DB 1; Length 358;  
 Best Local Similarity 64.3%; Pred. No. 5.8e-21;  
 Matches 45; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

QY 1 TLWGEFQVDSRGCGQTSNDAAEALNASSKEALQIREKIPKYLFOFHNLNSL 60  
 DB 109 TVMGQFQIDGSRGCGQTSNDAAEALNASSKEALQIREKIPKYLFOFHNLNSL 168  
 QY 61 DRIFDKTPE 70  
 DB 169 DRIFQEPAP 178

RESULT 4  
 ID VALL CLVW STANDARD; PRT: 358 AA.  
 AC P14972;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DE 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE ALI PROTEIN (40.4 KDA PROTEIN).  
 OS Cassava latent virus (strain Nigerian).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 RN ([1])  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-90174930; PubMed-2308831;  
 RA Morris B., Coates L., Love S., Richardson K., Eddy P.;  
 RT Nucleotide sequence of the infectious cloned DNA components of  
 RT African cassava mosaic virus (Nigerian strain).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))

CC EMBL: X17095; CAA34953.1; -  
 DR PIR: S07594; S07594.  
 DR INTERPRO: IPR001191;  
 DR PFAM: PF00749; Gemin\_A1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATCLV1.  
 KW ATP-binding.  
 FT NP\_BIND 220 227 ATP (POTENTIAL).  
 SQ SEQUENCE 358 AA: 40435 MW; 1DB16B80CB2D5E2C CRC64;

Query Match 69.0%; Score 251; DB 1; Length 358;  
 Best Local Similarity 64.3%; Pred. No. 5.8e-21;  
 Matches 45; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

QY 1 TLWGEFQVDSRGCGQTSNDAAEALNASSKEALQIREKIPKYLFOFHNLNSL 60  
 DB 109 TVMGQFQIDGSRGCGQTSNDAAEALNASSKEALQIREKIPKYLFOFHNLNSL 168

```

QY 61 DRIFTKPTPEP 70
DB 169 DRIFQEPFAP 178

RESULT 5
VALL_TYLCU
ID VALL_TYLCU STANDARD; PRT; 362 AA.
AC P36279; 1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE AL1 PROTEIN (C1 PROTEIN).
GN Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OS Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN SOURCE FROM N.A.
EX MEDLINE=93119776; PubMed=8423446;
HA DRY I.B., Riden J.E., Krake L.R., Mullineux P.M., Rezaian M.A.;
RT "Nucleotide sequence and genome organization of tomato leaf curl
geminivirus.";
RL J. Gen. Virol. 74:147-151(1993).
CC -1. SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL: S53251; CAS90888.1;
DR PIR: J01887; J01887;
DR INTERPRO: IP0001191;
DR PFAM: PF00799; Gemini_AL1; 1.
DR PRINTS: PR00227; GEMCOATALL.
DR PRINTS: PR00228; GEMCOATCLVL1.
KW ATP-binding. 221 228 ATP (POTENTIAL).
FT NP_BIND 362 AA; 41197 MW; 3436718/4704098 CRC64;
SQ SEQUENCE 362 AA; 41197 MW; 3436718/4704098 CRC64;

Query Match 67.0%; Score 244; DB 1; Length 362;
Best Local Similarity 56.5%; Pred.No. 3.6e-20;
Matches 48; Conservative 9; Mismatches 12; Indels 16; Gaps 1;

QY 1 TLVGEFQVDSRAGGCGTSDAAEAALNASSKEALQIREKIPKYLQFHNLSNL 60
DB 110 TLVGEFQVDSRAGGCGTSDAAEAALNASSKEALQIREKIPKYLQFHNLSNL 169
QY 61 DRI
DB 170 DRIFTPLVSPPLSSFDKPE 194

RESULT 6
VALL_TYLCU
ID VALL_TYLCU STANDARD; PRT; 359 AA.
AC P36509;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DE AL1 PROTEIN (C1 PROTEIN).
GN Tomato yellow leaf curl virus (strain Murcia) (TYLCV).
OS Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN SOURCE FROM N.A.
EX MEDLINE=94256836; PubMed=8198442;
RA Noris E., Hidalgo E., Accotto G., Moriones E.;
RT "High similarity among the tomato yellow leaf curl virus isolates
from the west Mediterranean basin: the nucleotide sequence of an
infectious clone from Spain";
RL Arch. Virol. 135:101-104(1994).
CC -1. SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL: 225751; CAA81026.1;
DR PIR: S39211; S39211;
DR INTERPRO: IP0001191;
DR PFAM: PF00799; Gemini_AL1; 1.
DR PRINTS: PR00227; GEMCOATALL.
DR PRINTS: PR00228; GEMCOATCLVL1.
KW ATP-binding. 321 228 ATP (POTENTIAL).
FT NP_BIND 359 AA; 41065 MW; 20170A51EF80A3EC CRC64;
SQ SEQUENCE 359 AA; 41065 MW; 20170A51EF80A3EC CRC64;

Query Match 65.1%; Score 237; DB 1; Length 359;
Best Local Similarity 62.3%; Pred.No. 2.4e-19;
Matches 43; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

QY 2 LMKGEFQVDSRAGGCGTSDAAEAALNASSKEALQIREKIPKYLQFHNLSNL 61
DB 111 LMKGEFQVDSRAGGCGTSDAAEAALNASSKEALQIREKIPKYLQFHNLSNL 170
QY 62 RLEKTPPEP 70
DB 171 RVFQVFPAP 179

RESULT 7
VALL_PHVU
ID VALL_PHVU STANDARD; PRT; 349 AA.
AC P36280;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE AL1 PROTEIN.
GN AL1.
OS Pepper huasteco virus (PHV).
OS Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN SOURCE FROM N.A.
EX MEDLINE=94015007; PubMed=8409944;
RA Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,
Rivera-Bustamante R.F.;
RT "Complete nucleotide sequence of pepper huasteco virus: analysis and
comparison with bipartite geminiviruses.";
RL Arch. Virol. 74:2225-2231(1995).
CC -1. SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL: X70418; CAA49856.1;
DR PIR: S31875; S31875;
DR PIR: JQ2300; JQ2300;
DR INTERPRO: IP0001191;

```

```

DR INTERPRO: IPR001301;
DR PFAM: PF00799; Gemin1_AL1; 1.
DR PRINTS: PR00227; GEMCOATALL.
DR KEGG: K01419; 221 228 ATP (BY SIMILARITY).
FT NP_BIND 221 228
SQ SEQUENCE 349 AA; 39722 MW; D5F4E76D56370F4 CRC64;

Query Match 64.04; Score 233; DB 1; Length 349;
Best Local Similarity 61.4%; Pred. No. 5.9e-19;
Matches 43; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

QY 1 TLVGEQVDGRSGGCGTSDNDAARALNASSKEALQIIRKIPKYLFOFHNLSNL 60
DB 110 TLVGEQVDGRSGGCGTSDNDAARALNASSKEALQIIRKIPKYLFOFHNLSNL 169
QY 61 DRIFKTPPEP 70
DB 170 NRIFKTPPEP 179

RESULT 8
ID VAL1_TYLCM STANDARD: PRT: 359 AA.
AC P27260; (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DR PRINTS: PR00227; GEMCOATALL.
DE ALL PROTEIN (C1 PROTEIN).
GN C1.
OS Tomato yellow leaf curl virus (strain Marmande) (TYLCV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP SEQUENCE FROM N.A.
RA Khoury-Pour A., Bendabmane M., Matzelt V., Accotto G.P., Crespi S.,
RA Gronenborn B.; PubMed:1840676;
RT "Tomato yellow leaf curl virus from Sardinia is a
RT whitefly-transmitted monopartite geminivirus.";
RL Nucleic Acids Res. 19:6763-6769(1991).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL: X61153; CAA43466.1;
DR PIR: S22593; S22593.
DR INTERPRO: IPR001191;
DR INTERPRO: IPR001301;
DR PFAM: PF00799; Gemin1_AL1; 1.
DR PRINTS: PR00227; GEMCOATALL.
DR KEGG: K01419; 220 227 ATP (POTENTIAL).
FT NP_BIND 220 227
SQ SEQUENCE 359 AA; 40733 MW; 9717B40A7C93EFA7 CRC64;

Query Match 63.24; Score 230; DB 1; Length 359;
Best Local Similarity 59.4%; Pred. No. 1.3e-18;
Matches 41; Conservative 11; Mismatches 17; Indels 0; Gaps 0;

QY 2 LVWGEQVDGRSGGCGTSDNDAARALNASSKEALQIIRKIPKYLFOFHNLSNL 61
DB 111 LVWGEQVDGRSGGCGTSDNDAARALNASSKEALQIIRKIPKYLFOFHNLSNL 170
QY 62 DRIFKTPPEP 70
DB 170 NRIFKTPPEP 179

RESULT 9
ID VAL1_ABMV STANDARD: PRT: 353 AA.
AC P05175;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DR PRINTS: PR00227; GEMCOATALL.
DE ALL PROTEIN (40.2 KDA PROTEIN).
GN AC1.
OS Bean golden mosaic virus
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP SEQUENCE FROM N.A.
RA Howarth A.J., Caton J., Bossert M., Goodman R.M.;
RA Nucleotide sequence of bean golden mosaic virus and a model for gene
RT regulation in geminiviruses.*;
RL Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL: M10070; AAA46318.1;
DR INTERPRO: IPR001191;
DR INTERPRO: IPR001301;
DR PFAM: PF00799; Gemin1_AL1; 1.
DR PRINTS: PR00227; GEMCOATALL.
DR KEGG: K01419; 222 229 ATP (POTENTIAL).
FT NP_BIND 222 229
SQ SEQUENCE 353 AA; 40190 MW; 80FA779DF6029A34 CRC64;

Query Match 62.64; Score 228; DB 1; Length 353;
Best Local Similarity 61.4%; Pred. No. 2.2e-18;
Matches 43; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

QY 1 TLVGEQVDGRSGGCGTSDNDAARALNASSKEALQIIRKIPKYLFOFHNLSNL 60
DB 110 TLVGEQVDGRSGGCGTSDNDAARALNASSKEALQIIRKIPKYLFOFHNLSNL 169
QY 61 DRIFKTPPEP 70
DB 170 NRIFKTPPEP 179

RESULT 10
ID VAL1_ABMV STANDARD: PRT: 355 AA.
AC P21947; 1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DR PRINTS: PR00227; GEMCOATALL.
DE ALL PROTEIN.
GN AC1.
OS Abutilon mosaic virus (isolate West India).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP SEQUENCE FROM N.A.
RA Frieschmuth T., Zimmatt G., Jeske H.;
RA "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic
RT as well as eukaryotic features.*";
RL Virology 178:461-468(1990).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.

```





OS Tomato yellow leaf curl virus (TYLCV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 RN [1] JENCE FROM N.A.  
 RX MEDLINE-92024070; PubMed-1926771;  
 RA Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;  
 RT "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus  
 with a single genomic component.";  
 RL Virology 185:151-161(1991).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; X15556; C0433688.1; -.  
 DR PIR; X15556; Q0CYS1; -.  
 DR INTERPRO: IPR001191; -.  
 DR PFAM: PF00799; Gemini\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATL1.  
 KW ATP-binding; 319 226 ATP (POTENTIAL).  
 FT NP-BIND 319 226  
 SQ SEQUENCE 357 AA; 40678 MW; 939A068E1A392A7 CRC64;

Query Match 59.1%; Score 215; DB 1; Length 357;  
 Best Local Similarity 67.24; Pred. No. 6.4e-17;  
 Matches 41; Conservative 10; Mismatches 10; Indels 0; Gaps 0;  
 QY 4 NGCFVQVGRSGCCGTSNDAAEALNASKKEALQIREKIPKYLQFHNLSNLDRI 63  
 DB 111 FGVSQIDCHSANGSGQSDANDAAEALNASKKEALNILKPKAPDYILOFHNLSNLDRI 170  
 QY 64 F 64  
 DB 171 F 171

RESULT 14  
 VAL1\_SICV STANDARD; PRT; 347 AA.  
 AC P29048;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DE ALL PROTEIN  
 OS Squash leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-91082449; PubMed-1984669;  
 RA Lazarowitz S.G., Lazdins I.B.;  
 RT "Identification of a complete nucleotide sequence of the cloned genomic  
 fragment of a tobacco etch virus geminivirus with a broad  
 host range phenotype.";  
 RL Virology 180:58-69(1991).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; M36183; AAC32410.1; ALT-INIT.

DR PIR; C36785; Q0CYS1.  
 DR INTERPRO: IPR001191; -.  
 DR PFAM: PF00799; Gemini\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATL1.  
 KW ATP-binding; 218 225  
 FT NP-BIND 218 225 ATP (POTENTIAL).  
 SQ SEQUENCE 347 AA; 39110 MW; AFDABDE122110E CRC64;

Query Match 34.1%; Score 124; DB 1; Length 347;  
 Best Local Similarity 37.9%; Pred. No. 1e-06; 23; Indels 1;  
 Matches 25; Conservative 14; Mismatches 4; Gaps 1;  
 QY 5 GFQFVQVGRSGCCGTSNDAAEALNASKKEALQIREKIPKYLQFHNLSNLDRI 64  
 DB 116 QQTKYKSG-----GSKSNKODVYHNAVNGSAGEALDTIKAGDPKFTFTVYVHNLINVERLF 171  
 QY 65 DKTPPP 70  
 DB 172 OKTPPP 177

RESULT 15  
 Y175\_HELRY STANDARD; PRT; 299 AA.  
 ID Y175\_HELRY  
 DC Q22A07; 2000 (Rel. 40, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE HYPOTHETICAL PROTEIN HP0175 PRECURSOR.  
 GN HP0175 OR JHP0161.  
 OS Helicobacter pylori J99 (Campylobacter pylori J99).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-99120557; PubMed-9923682;  
 RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.B., Doig P.C.,  
 RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,  
 RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,  
 RA Gibson R., Merberg D., Mills S.B., Jiang Q., Taylor D.E., Novis G.F.,  
 RT "First J sequence comparison of two unrelated isolates of the human  
 gastric pathogen Helicobacter pylori.";  
 RL Nature 397:176-180(1999).  
 CC -1- SIMILARITY: BELONGS TO THE PPIC/PARVULIN FAMILY OF ROTASES.  
 CC -1- STRONG, TO C.JEJUNI CBF2.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; AE00145; J04005744.1; -.  
 DR PIR; AE00145; J04005744.1; -.  
 DR INTERPRO: IPR000297; -.  
 DR PFAM: PF00639; Rotamase; 1.  
 DR POSITIVE; PS01096; PPIC-PPPIASE; 1.  
 KW Hypothetical protein; Isomerase; Rotamase; Signal.  
 FT SIGNAL 1 21  
 FT CHAIN 12 299  
 FT REGION 12 299 POTENTIAL.  
 FT REGION 12 299 HYPOTHETICAL PROTEIN HP0175.  
 SQ SEQUENCE 299 AA; 34040 MW; 9C037B1CD1110143 CRC64;

Query Match 17.7%; Score 64.5; DB 1; Length 299;  
 Best Local Similarity 28.1%; Pred. No. 4.1;  
 Matches 18; Conservative 14; Mismatches 11; Indels 21; Gaps 2;

Qy 22 NDAADALNAS-----SKBEALQIREKIEKYLFOFHNLSLDR 62  
Db 93 NERAKENQTPPEKAMMEAVKKQALVEFWAKQAEVKKIOIPEREMODFY--NARKDO 150  
Qy 63 IFDK 66  
Db 151 LFVK 154

Search completed: February 3, 2001, 02:24:24  
Job time: 636 sec

---

**THIS PAGE BLANK (USPTO)**

1

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: February 3, 2001, 01:40:48 ; Search time 118.74 Seconds  
(without alignments)  
40.029 Million cell updates/sec

Title: US-09-289-346a-1  
Perfect score: 364  
Sequence: 1 TLVWGFQVQDGRSARGSCQT.....QFPHNLSNLDIFQDTPK 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 66: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	364	100.0	352	1 QOCVLI	Al1 protein - toma
2	264	72.5	361	1 QOCVPT	Al1 protein - toma
3	251	69.0	358	2 S07594	hypothetical prote
4	244	67.0	362	1 JQ1887	Al1 protein - toma
5	237	65.1	359	2 S92211	gene C1 protein -
6	233	64.0	349	2 J03200	replicase - pepper
7	233	64.0	359	2 S1365	Al1 protein - pepp
8	227	62.4	358	1 JQ1870	hypothetical prote
9	227	62.4	351	2 J02237	Al1 protein - toma
10	227	62.4	355	1 QOCVM1	AV1 protein - abut
11	224	61.5	358	1 JQ1870	Al1 protein - toma
12	224	61.5	359	2 S92235	gene C1 protein -
13	224	61.5	385	2 S28360	Al1 protein - best
14	217	59.6	360	2 S59885	replication-associ
15	215	58.0	360	1 QOCVLI	Al1 protein - toma
16	215	58.0	337	1 QOCVPT	Al1 protein - toma
17	124	34.1	347	1 QOCVSI	Al1 protein - toma
18	68	18.7	587	2 JC1419	PC gamma (Igc) roc
19	64.5	17.7	299	2 B71967	probable peptidyl-
20	64.5	17.7	1713	2 A55347	adhesive ligand ep
21	63.5	17.4	1610	2 A46227	voltage-dependent
22	63.5	17.4	1646	2 JH0422	voltage-dependent
23	63.5	17.4	2181	2 A49398	calcium channel al
24	63.5	17.4	2181	2 A49398	voltage-dependent
25	63.5	17.4	2203	2 T42742	voltage-dependent
26	62.5	17.2	447	2 T12544	hypothetical prote
27	62	17.0	295	2 D42452	Cl protein - tobac
28	61.5	16.9	481	2 T70091	probable phosphos
29	61	16.8	840	2 T36175	probable large ATP

30 61 16.8 1502 1 RGBYH1  
31 60.5 16.6 354 2 A75087  
32 60.5 16.6 1229 2 T48959  
33 60 16.5 447 2 S2437  
34 59.5 16.5 2181 2 A49398  
35 59 16.2 160 2 G82060  
36 59 16.2 316 2 G82085  
37 59 16.2 397 2 B71078  
38 59 16.2 416 2 A82892  
39 58.5 16.1 136 2 T22240  
40 58.5 16.1 297 1 MIVNRV  
41 58.5 16.1 599 2 G64541  
42 58.5 16.1 1044 2 T43000  
43 58.5 16.1 1044 2 T43000  
44 58.5 16.1 1265 2 T47626  
45 58.5 16.1 1743 2 T15893

CYCL1/CYC3 transcri  
acetyl ornithine d  
kinesin-like prote  
cup-diacylglycerol  
hypothetical prote  
hypothetical prote  
conserved hypotet  
probable NADH oxid  
hypothetical prote  
hypothetical prote  
nonstructural prote  
cell binding facto  
protein kinase p  
protein kinase bub  
structural mainten  
hypothetical prote

## ALIGNMENTS

RESULT 1  
QOCVLI  
C:Species: tomato golden mosaic virus  
A:Note: host Nicotiana sp. (tobacco)  
C:Date: 28-Aug-1995 #sequence\_revision 28-Aug-1985 #text\_change 08-Apr-1994  
E:Accession: A04170 Steink, V.E.; Coutts, R.H.A.; Buck, K.W.  
E:MO: 1.3.1 2157-2205, 1984  
A:Title: Complete nucleotide sequence of the infectious cloned DNA components of toma  
A:Reference number: A04163  
A:Accession: A04170  
A:Molecule type: DNA  
A:Residues: 1-352 <HAM>  
C:Comment: The genome consists of two circular, single-stranded DNA components, DNA A  
C:Genetics:  
A:Map position: segment A  
C:Superfamily: tomato golden mosaic virus Al1 protein

Query Match 100.0% Score 364; DB 1; Length 352;  
Best Local Similarity 100.0% Prod. NO. 2.46-33;  
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 TLVWGFQVQDGRSARGSCQTSDAAALNASSKEALQIIRKTPKYLQFPHNLSNLI 60  
|||||  
Db 111 TLVWGFQVQDGRSARGSCQTSDAAALNASSKEALQIIRKTPKYLQFPHNLSNLI 170  
Oy 61 DRIFDXTPEP 70  
|||||  
Db 171 DRIFDXTPEP 180

RESULT 2  
QOCVPT  
Al1 protein - potato yellow mosaic virus (isolate Venezuela)  
C:Species: potato yellow mosaic virus  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Jun-2000  
E:Accession: J00364  
E:MO: 1.3.1 1515-1520, 1991  
R:Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.  
J. Gen. Virol. 72, 1515-1520, 1991  
A:Title: The nucleotide sequence of the infectious cloned DNA components of potato ye  
A:Reference number: J00362; MUTD:91311403  
A:Accession: J00364  
A:Status: translation not shown  
A:Residues: 361 <COU>  
A:Cross-references: GB:D00940; NID:g222458; PIDN:BAA00782.1; PID:g222459  
C:Genetics:  
A:Map position: segment A  
C:Superfamily: tomato golden mosaic virus Al1 protein

Query Match  
Best Local Similarity 72.5%; Score 264; DB 1; Length 361;  
Matches 49; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

QY 1 TLWGEFQVDRSGRGCCQTSNDAAAEALNASSKEALQIIRKIPKYLFOFHNLNSL 60  
1 : |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: ||

Db 110 TLWGEFQVDRSGRGCCQTSNDAAAEALNASSKEALQIIRKIPKYLFOFHNLNSL 169  
1 : |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: ||

QY 61 DRIFKTPPE 69  
Db 170 DRIFKTPPE 178

RESULT 3  
S07594  
hypochemical protein, 40.4K - cassava latent virus (Nigerian isolate)  
C:Species: cassava latent virus  
C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 20-Sep-1999  
C:Accession: S07594  
R:Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.  
Nucleic Acids. Res. 18, 197-198, 1990  
A:Title: Nucleotide sequence of the infectious cloned DNA components of African cassava  
A:Reference number: S07590; MUID:90174930  
A:Accession: S07594  
A:Status: translation not shown  
A:Residues: 1-358 <D>  
A:Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376  
C:Genetics:  
A:Map position: segment DN41  
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 69.0%; Score 251; DB 2; Length 358;  
Best Local Similarity 64.3%; Pred. No. 1,2e-20;  
Matches 45; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

QY 1 TLWGEFQVDRSGRGCCQTSNDAAAEALNASSKEALQIIRKIPKYLFOFHNLNSL 60  
1 : |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: ||

Db 109 TLWGEFQVDRSGRGCCQTSNDAAAEALNASSKEALQIIRKIPKYLFOFHNLNSL 168  
1 : |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: ||

QY 61 DRIFKTPPE 70  
Db 169 DRIFKTPPE 178

RESULT 4  
S07594  
AL1 protein - tomato yellow leaf curl virus (strain Australlia)  
N:Alternate names: Cl protein  
C:Species: tomato yellow leaf curl virus  
C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 07-May-1999  
C:Accession: JQ1887  
R:Dry, I.B.; Rigden, J.B.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.  
J. Gen. Virol. 74, 147-151, 1993  
A:Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.  
A:Reference number: JQ1887; MUID:93139778  
A:Accession: JQ1887  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-362 <D>  
A:Cross-references: GB:S53251  
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 67.0%; Score 244; DB 1; Length 362;  
Best Local Similarity 56.5%; Pred. No. 7.5e-20;  
Matches 48; Conservative 9; Mismatches 12; Indels 16; Gaps 1;

QY 1 TLWGEFQVDRSGRGCCQTSNDAAAEALNASSKEALQIIRKIPKYLFOFHNLNSL 60  
1 : |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: ||

Db 110 TLWGEFQVDRSGRGCCQTSNDAAAEALNASSKEALQIIRKIPKYLFOFHNLNSL 169  
1 : |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: ||

QY 61 DRI-----FDKTPPE 69  
Db 170 DRIFTPLEVYVSFLSSSDFAVPE 194  
1 : |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: ||

RESULT 5  
S39211  
gene C1 protein - tomato yellow leaf curl virus  
C:Species: tomato yellow leaf curl virus  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 20-Sep-1999  
R:Morris, B.; Hidalgo, E.; Accotto, G.; Moriones, E.  
submitted to the EMBL Data Library, August 1993  
A:Description: High similarity among the tomato yellow leaf curl virus isolates from  
A:Reference number: S39209  
A:Accession: S39211  
A:Status: preliminary  
A:Residues: 1-359 <D>  
A:Cross-references: EMBL:Z25751; NID:g433655; PIDN:CAA81026.1; PID:g433658  
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 65.1%; Score 237; DB 2; Length 359;  
Best Local Similarity 62.3%; Pred. No. 4.5e-19;  
Matches 43; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

QY 2 LWGEFQVDRSGRGCCQTSNDAAAEALNASSKEALQIIRKIPKYLFOFHNLNSL 61  
1 : |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: ||

Db 111 LWGEFQVDRSGRGCCQTSNDAAAEALNASSKEALQIIRKIPKYLFOFHNLNSL 170  
1 : |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: ||

QY 62 RIFDKTPPE 70  
Db 171 RIFQVPPAP 179

RESULT 6  
JQ2300  
replicase - pepper huasteco virus (component A)  
N:Alternate names: OR AL1 protein  
C:Species: pepper huasteco virus  
C:Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 20-Sep-1999  
C:Accession: JQ2300  
R:Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante,  
J. Gen. Virol. 74, 2225-2231, 1993  
A:Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b  
A:Reference number: JQ2299; MUID:94015007  
A:Accession: JQ2300  
A:Residues: 1-349 <D>  
A:Cross-references: GB:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025  
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 64.0%; Score 233; DB 2; Length 349;  
Best Local Similarity 61.4%; Pred. No. 1.2e-16;  
Matches 43; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

QY 1 TLWGEFQVDRSGRGCCQTSNDAAAEALNASSKEALQIIRKIPKYLFOFHNLNSL 60  
1 : |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: ||

Db 110 TLWGEFQVDRSGRGCCQTSNDAAAEALNASSKEALQIIRKIPKYLFOFHNLNSL 169  
1 : |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: ||

QY 61 DRIFKTPPE 70  
Db 170 DRIFKTPPE 179

RESULT 7  
S31875  
AL1 protein - pepper rizado amarillo virus  
C:Species: pepper rizado amarillo virus



QY 1 TLWGEFVDGSRGCGCOTSDAAAPALNASSKEALQIIREKIPKYLFOFPHLNSNL 60  
 Db 107 TLWGEFVDGSRGCGCOTSDAAAPALNASSKEALQIIREKIPKYLFOFPHLNSNL 166  
 QY 61 DRIFDKTPEP 70  
 Db 167 BRIFAKAPEP 176

RESULT 12  
 S39235  
 S39235  
 S39235  
 gene C1 protein - tomato yellow leaf curl virus  
 C:Species: tomato yellow leaf curl virus  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 08-Sep-1997  
 C:Accession: S39235  
 R:Crespi, S.; Norris, E.; Vaira, A.; Bosco, D.; Accotto, G.  
 submitted to the EMBL data library, December 1993  
 A:Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity.  
 A:Accession: S39235  
 A:Accession: S39235  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-359 <R>  
 A:Cross-references: EMBL:Z28390; NID:g1041671; PID:g1334964  
 C:Superfamily: tomato golden mosaic virus ALI protein

Query Match 61.5%; Score 224; DB 2; Length 359;  
 Best Local Similarity 58.0%; Pred. No. 1.3e-17;  
 Matches 40; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

QY 2 LWGEFVDGSRGCGCOTSDAAAPALNASSKEALQIIREKIPKYLFOFPHLNSNL 61  
 Db 111 LWGEFVDGSRGCGCOTSDAAAPALNASSKEALQIIREKIPKYLFOFPHLNSNL 170  
 QY 62 DRIFDKTPEP 70  
 Db 171 KVFQVPPAP 179

RESULT 13  
 S28360  
 S28360  
 ALI protein - beet curly top virus  
 C:Species: beet curly top virus  
 C:Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 20-Sep-1999  
 C:Accession: S28360  
 R:Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.  
 A:Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly top virus  
 A:Reference number: S28360  
 A:Accession: S28360  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-385 <STA>  
 A:Cross-references: GB:M24597; EMBL:X04144; NID:g210678; PID:AAA42751.1; PID:g210679  
 C:Superfamily: tomato golden mosaic virus ALI protein

Query Match 61.5%; Score 224; DB 2; Length 385;  
 Best Local Similarity 57.1%; Pred. No. 1.4e-17;  
 Matches 40; Conservative 16; Mismatches 14; Indels 0; Gaps 0;

QY 1 TLWGEFVDGSRGCGCOTSDAAAPALNASSKEALQIIREKIPKYLFOFPHLNSNL 60  
 Db 137 TLWGEFVDGSRGCGCOTSDAAAPALNASSKEALQIIREKIPKYLFOFPHLNSNL 196  
 QY 61 DRIFDKTPEP 70  
 Db 197 OKIFQRPDP 206

RESULT 14

S59885  
 replication-associated protein C1 - tomato yellow leaf curl virus  
 C:Species: tomato yellow leaf curl virus  
 C:Date: 14-Jul-1996 #sequence\_revision 01-Mar-1996 #text\_change 20-Sep-1999  
 C:Accession: S59885  
 R:Hong, Y.; Harrison, B.D.  
 submitted to the EMBL data library, February 1995  
 A:Description: Nucleotide sequences from tomato leaf curl viruses from different countries  
 A:Accession: S59885  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-360 <KON>  
 A:Cross-references: EMBL:Z48182; NID:g444838; PID:CAA88229.1; PID:g974211  
 C:Superfamily: tomato golden mosaic virus ALI protein

Query Match 59.6%; Score 217; DB 2; Length 360;  
 Best Local Similarity 62.1%; Pred. No. 8.1e-17;  
 Matches 41; Conservative 12; Mismatches 13; Indels 0; Gaps 0;

QY 4 WGEFVDGSRGCGCOTSDAAAPALNASSKEALQIIREKIPKYLFOFPHLNSNL 63  
 Db 113 FGWFDGSRGCGCOTSDAAAPALNASSKEALQIIREKIPKYLFOFPHLNSNL 172  
 QY 64 FDKTPE 69  
 Db 173 FTFSAE 178

RESULT 15  
 QQCVC1  
 ALI protein - tomato yellow leaf curl virus  
 A:Alternate names: C1 protein curl virus  
 A:Accession: S28360  
 C:Date: 30-Jun-1993 #sequence\_revision 30-Jun-1992 #text\_change 23-Jul-1999  
 C:Accession: D40779  
 R:Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.  
 Virology 185, 151-161, 1991  
 A:Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin  
 A:Reference number: A40779; MUID:92024070  
 A:Accession: S28360  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-357 <NAV>  
 A:Cross-references: GB:X15656; NID:g62204; PID:CAA33688.1; PID:g62207  
 C:Superfamily: tomato golden mosaic virus ALI protein

Query Match 59.1%; Score 215; DB 1; Length 357;  
 Best Local Similarity 67.2%; Pred. No. 1.3e-16;  
 Matches 41; Conservative 10; Mismatches 10; Indels 0; Gaps 0;

QY 4 WGEFVDGSRGCGCOTSDAAAPALNASSKEALQIIREKIPKYLFOFPHLNSNL 63  
 Db 111 FGWFDGSRGCGCOTSDAAAPALNASSKEALQIIREKIPKYLFOFPHLNSNL 170  
 QY 64 F 64  
 Db 171 F 171

Search completed: February 3, 2001, 02:17:36  
 Job time: 2208 sec



1	230	63.2	359	3	US-08-809-1038-2	Sequence 2, Appl
2	230	63.2	359	3	US-08-809-1038-4	Sequence 4, Appl
3	230	63.2	359	3	US-08-809-1038-6	Sequence 6, Appl
4	230	63.2	359	3	US-08-809-1038-8	Sequence 8, Appl
5	64.5	17.7	1713	3	US-08-600-982-4	Sequence 24, Appl
6	64.5	17.7	1713	3	US-08-600-982-24	Sequence 24, Appl
7	64.5	17.7	2163	4	US-07-954-52621-24	Sequence 24, Appl
8	63.5	17.4	2161	1	US-08-455-543A-49	Sequence 49, Appl
9	63.5	17.4	2161	1	US-08-455-543A-51	Sequence 51, Appl
10	63.5	17.4	2161	2	US-08-233-305C-49	Sequence 49, Appl
11	63.5	17.4	2161	2	US-08-233-305C-51	Sequence 51, Appl
12	63.6	16.5	2161	2	US-08-672-863-11	Sequence 11, Appl
13	64.5	17.7	2163	4	US-07-954-52621-24	Sequence 24, Appl
14	58	15.9	844	3	US-09-009-267-20	Sequence 20, Appl
15	57.5	15.8	439	1	US-07-637-870-9	Sequence 9, Appl
16	57.5	15.8	439	1	US-07-637-399-6	Sequence 6, Appl
17	57.5	15.8	439	1	US-07-637-399-6	Sequence 6, Appl
18	56.5	15.5	334	3	US-08-112-703-3	Sequence 3, Appl
19	56.5	15.5	334	3	US-08-112-703-6	Sequence 6, Appl
20	56.5	15.5	334	3	US-08-112-703-11	Sequence 11, Appl
21	56	15.4	602	2	US-07-419-653-6	Sequence 6, Appl
22	56	15.4	771	1	US-07-923-976-5	Sequence 5, Appl
23	56	15.4	783	5	5422248-2	Patent No. 5422248
24	56	15.4	836	1	US-07-923-976-4	Sequence 4, Appl
25	56	15.4	863	1	US-07-923-976-8	Sequence 8, Appl
26	54	14.8	434	3	US-08-446-100-30	Sequence 30, Appl
27	53	14.6	456	3	US-08-446-100-50	Sequence 50, Appl
28	53	14.6	456	3	US-08-446-100-26	Sequence 26, Appl
29	53	14.6	456	3	US-08-446-100-27	Sequence 27, Appl
30	53	14.6	456	3	US-08-446-100-27	Sequence 27, Appl

Query Match 53.2%; Score 230; DB 3; Length 359;  
Best Local Similarity 59.4%; Pred. No. 1.3e-22;  
Matches 41: Conservative 11; Mismatches 17; Indels

```

QY 2 LVNGFQVDRSARGCQTSNDAAEALNASSKEEAQLIIREKIPKYLQFHHNSNLD 61
Db 111 LVNGFQVDRSARGCQTSNDAAEALNASSKEEAQLIIREKIPKYLQFHHNSNLD 170
QY 62 RIFDKTPEP 70
Db 171 KVEQVPPAP 179

RESULT 2 103B-4
US-08-809-103B-4
Patent No. 6133505
GENERAL INFORMATION:
APPLICANT: GRONENBORN, Bruno
TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: YOUNG & THOMPSON
STREET: 745 South 23rd Street
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 17-MAR-1997
CLASSIFICATION: 800
PRIORITY INFORMATION:
PRIOR APPLICATION NUMBER: US/08/809,103B
FILING DATE: 17-MAR-1997
APPLICATION NUMBER: FR 94.11040
PRIORITY INFORMATION:
PCT NUMBER: WO PCT/FR95/01192
ATTORNEY/AGENT INFORMATION:
NAME: PATCH, Andrew J.
REGISTRATION NUMBER: 32,925
REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 521-2297
TELEX: 248425 EMBON
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 359 amino acids
TYPE: amino acid
MOLECULE TYPE: linear
US-08-809-103B-4

Query Match 63.2%; Score 230; DB 3; Length 359;
Best Local Similarity 59.4%; Pred. No. 1.3e-22;
Matches 41; Conservative 11; Mismatches 17; Indels 0;

QY 2 LVNGFQVDRSARGCQTSNDAAEALNASSKEEAQLIIREKIPKYLQFHHNSNLD 61
Db 111 LVNGFQVDRSARGCQTSNDAAEALNASSKEEAQLIIREKIPKYLQFHHNSNLD 170
QY 62 RIFDKTPEP 70
Db 171 KVEQVPPAP 179

RESULT 3
US-08-809-103B-6
Patent No. 6133505
GENERAL INFORMATION:
APPLICANT: GRONENBORN, Bruno
TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: YOUNG & THOMPSON
STREET: 745 South 23rd Street
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 17-MAR-1997
CLASSIFICATION: 800
PRIORITY INFORMATION:
PRIOR APPLICATION NUMBER: US/08/809,103B
FILING DATE: 17-MAR-1997
APPLICATION NUMBER: FR 94.11040
PRIORITY INFORMATION:
PCT NUMBER: WO PCT/FR95/01192
ATTORNEY/AGENT INFORMATION:
NAME: PATCH, Andrew J.
REGISTRATION NUMBER: 32,925
REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 521-2297
TELEX: 248425 EMBON
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 359 amino acids
TYPE: amino acid
MOLECULE TYPE: linear
US-08-809-103B-6

Query Match 63.2%; Score 230; DB 3; Length 359;
Best Local Similarity 59.4%; Pred. No. 1.3e-22;
Matches 41; Conservative 11; Mismatches 17; Indels 0;

QY 2 LVNGFQVDRSARGCQTSNDAAEALNASSKEEAQLIIREKIPKYLQFHHNSNLD 61
Db 111 LVNGFQVDRSARGCQTSNDAAEALNASSKEEAQLIIREKIPKYLQFHHNSNLD 170
QY 62 RIFDKTPEP 70
Db 171 KVEQVPPAP 179

RESULT 4
US-08-809-103B-8
Patent No. 6133505
GENERAL INFORMATION:
APPLICANT: GRONENBORN, Bruno
TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: YOUNG & THOMPSON
STREET: 745 South 23rd Street
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 17-MAR-1997
CLASSIFICATION: 800
PRIORITY INFORMATION:
PRIOR APPLICATION NUMBER: US/08/809,103B
FILING DATE: 17-MAR-1997
APPLICATION NUMBER: FR 94.11040
PRIORITY INFORMATION:
PCT NUMBER: WO PCT/FR95/01192
ATTORNEY/AGENT INFORMATION:
NAME: PATCH, Andrew J.
REGISTRATION NUMBER: 32,925
REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 521-2297
TELEX: 248425 EMBON
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 359 amino acids
TYPE: amino acid
MOLECULE TYPE: linear
US-08-809-103B-8

Query Match 63.2%; Score 230; DB 3; Length 359;
Best Local Similarity 59.4%; Pred. No. 1.3e-22;
Matches 41; Conservative 11; Mismatches 17; Indels 0; Gaps 0;

QY 2 LVNGFQVDRSARGCQTSNDAAEALNASSKEEAQLIIREKIPKYLQFHHNSNLD 61
Db 111 LVNGFQVDRSARGCQTSNDAAEALNASSKEEAQLIIREKIPKYLQFHHNSNLD 170
QY 62 RIFDKTPEP 70
Db 171 KVEQVPPAP 179

```

```

; Sequence 6, Application US/08809103B
; Patent No. 6133505
; GENERAL INFORMATION:
; APPLICANT: GRONENBORN, Bruno
; TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: YOUNG & THOMPSON
; STREET: 745 South 23rd Street
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 17-MAR-1997
; CLASSIFICATION: 800
; PRIORITY INFORMATION:
; PRIOR APPLICATION NUMBER: FR 94.11040
; FILING DATE: 15-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR95/01192
; PRIORITY INFORMATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: PATCH, Andrew J.
; REGISTRATION NUMBER: 32,925
; REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 521-2297
; TELEX: 248425 EMBON
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 359 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-809-103B-6

Query Match 63.2%; Score 230; DB 3; Length 359;
Best Local Similarity 59.4%; Pred. No. 1.3e-22;
Matches 41; Conservative 11; Mismatches 17; Indels 0; Gaps 0;

QY 2 LVNGFQVDRSARGCQTSNDAAEALNASSKEEAQLIIREKIPKYLQFHHNSNLD 61
Db 111 LVNGFQVDRSARGCQTSNDAAEALNASSKEEAQLIIREKIPKYLQFHHNSNLD 170
QY 62 RIFDKTPEP 70
Db 171 KVEQVPPAP 179

RESULT 4
US-08-809-103B-8
Patent No. 6133505
GENERAL INFORMATION:
APPLICANT: GRONENBORN, Bruno
TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: YOUNG & THOMPSON
STREET: 745 South 23rd Street
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 17-MAR-1997
CLASSIFICATION: 800
PRIORITY INFORMATION:
PRIOR APPLICATION NUMBER: FR 94.11040
FILING DATE: 15-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR95/01192
PRIORITY INFORMATION:
ATTORNEY/AGENT INFORMATION:
NAME: PATCH, Andrew J.
REGISTRATION NUMBER: 32,925
REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 521-2297
TELEX: 248425 EMBON
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 359 amino acids
TYPE: amino acid
MOLECULE TYPE: linear
US-08-809-103B-8

Query Match 63.2%; Score 230; DB 3; Length 359;
Best Local Similarity 59.4%; Pred. No. 1.3e-22;
Matches 41; Conservative 11; Mismatches 17; Indels 0; Gaps 0;

QY 2 LVNGFQVDRSARGCQTSNDAAEALNASSKEEAQLIIREKIPKYLQFHHNSNLD 61
Db 111 LVNGFQVDRSARGCQTSNDAAEALNASSKEEAQLIIREKIPKYLQFHHNSNLD 170
QY 62 RIFDKTPEP 70
Db 171 KVEQVPPAP 179

```

```

1  COUNTRY: U.S.A.
2  ZIP: 22202
3  COMPUTER READABLE FORM:
4  NAME: Shelton, Dennis K.
5  REGISTRATION NUMBER: 26,997
6  TOPLOGY: linear
7  OPERATING SYSTEM: PC-DOS/MS-DOS
8  SOFTWARE: Patent Release #1.0, Version #1.30
9  CURRENT APPLICATION DATA:
10 FILING DATE: 17-MAR-1997
11 FILING DATE: 15-SEP-1995
12 FILING DATE: 15-SEP-1995
13 FILING DATE: 15-SEP-1995
14 FILING DATE: 15-SEP-1995
15 FILING DATE: 15-SEP-1995
16 FILING DATE: 15-SEP-1995
17 FILING DATE: 15-SEP-1995
18 FILING DATE: 15-SEP-1995
19 FILING DATE: 15-SEP-1995
20 FILING DATE: 15-SEP-1995
21 FILING DATE: 15-SEP-1995
22 FILING DATE: 15-SEP-1995
23 FILING DATE: 15-SEP-1995
24 FILING DATE: 15-SEP-1995
25 FILING DATE: 15-SEP-1995
26 FILING DATE: 15-SEP-1995
27 FILING DATE: 15-SEP-1995
28 FILING DATE: 15-SEP-1995
29 FILING DATE: 15-SEP-1995
30 FILING DATE: 15-SEP-1995
31 FILING DATE: 15-SEP-1995
32 FILING DATE: 15-SEP-1995
33 FILING DATE: 15-SEP-1995
34 FILING DATE: 15-SEP-1995
35 FILING DATE: 15-SEP-1995
36 FILING DATE: 15-SEP-1995
37 FILING DATE: 15-SEP-1995
38 FILING DATE: 15-SEP-1995
39 FILING DATE: 15-SEP-1995
40 FILING DATE: 15-SEP-1995
41 FILING DATE: 15-SEP-1995
42 FILING DATE: 15-SEP-1995
43 FILING DATE: 15-SEP-1995
44 FILING DATE: 15-SEP-1995
45 FILING DATE: 15-SEP-1995
46 FILING DATE: 15-SEP-1995
47 FILING DATE: 15-SEP-1995
48 FILING DATE: 15-SEP-1995
49 FILING DATE: 15-SEP-1995
50 FILING DATE: 15-SEP-1995
51 FILING DATE: 15-SEP-1995
52 FILING DATE: 15-SEP-1995
53 FILING DATE: 15-SEP-1995
54 FILING DATE: 15-SEP-1995
55 FILING DATE: 15-SEP-1995
56 FILING DATE: 15-SEP-1995
57 FILING DATE: 15-SEP-1995
58 FILING DATE: 15-SEP-1995
59 FILING DATE: 15-SEP-1995
60 FILING DATE: 15-SEP-1995
61 FILING DATE: 15-SEP-1995
62 FILING DATE: 15-SEP-1995
63 FILING DATE: 15-SEP-1995
64 FILING DATE: 15-SEP-1995
65 FILING DATE: 15-SEP-1995
66 FILING DATE: 15-SEP-1995
67 FILING DATE: 15-SEP-1995
68 FILING DATE: 15-SEP-1995
69 FILING DATE: 15-SEP-1995
70 FILING DATE: 15-SEP-1995
71 FILING DATE: 15-SEP-1995
72 FILING DATE: 15-SEP-1995
73 FILING DATE: 15-SEP-1995
74 FILING DATE: 15-SEP-1995
75 FILING DATE: 15-SEP-1995
76 FILING DATE: 15-SEP-1995
77 FILING DATE: 15-SEP-1995
78 FILING DATE: 15-SEP-1995
79 FILING DATE: 15-SEP-1995
80 FILING DATE: 15-SEP-1995
81 FILING DATE: 15-SEP-1995
82 FILING DATE: 15-SEP-1995
83 FILING DATE: 15-SEP-1995
84 FILING DATE: 15-SEP-1995
85 FILING DATE: 15-SEP-1995
86 FILING DATE: 15-SEP-1995
87 FILING DATE: 15-SEP-1995
88 FILING DATE: 15-SEP-1995
89 FILING DATE: 15-SEP-1995
90 FILING DATE: 15-SEP-1995
91 FILING DATE: 15-SEP-1995
92 FILING DATE: 15-SEP-1995
93 FILING DATE: 15-SEP-1995
94 FILING DATE: 15-SEP-1995
95 FILING DATE: 15-SEP-1995
96 FILING DATE: 15-SEP-1995
97 FILING DATE: 15-SEP-1995
98 FILING DATE: 15-SEP-1995
99 FILING DATE: 15-SEP-1995
100 FILING DATE: 15-SEP-1995

```

```

1  CLASSIFICATION: 435
2  ATTORNEY/AGENT INFORMATION:
3  NAME: Shelton, Dennis K.
4  REGISTRATION NUMBER: 26,997
5  TOPLOGY: linear
6  OPERATING SYSTEM: PC-DOS/MS-DOS
7  SOFTWARE: Patent Release #1.0, Version #1.30
8  CURRENT APPLICATION DATA:
9  FILING DATE: 17-MAR-1997
10 FILING DATE: 15-SEP-1995
11 FILING DATE: 15-SEP-1995
12 FILING DATE: 15-SEP-1995
13 FILING DATE: 15-SEP-1995
14 FILING DATE: 15-SEP-1995
15 FILING DATE: 15-SEP-1995
16 FILING DATE: 15-SEP-1995
17 FILING DATE: 15-SEP-1995
18 FILING DATE: 15-SEP-1995
19 FILING DATE: 15-SEP-1995
20 FILING DATE: 15-SEP-1995
21 FILING DATE: 15-SEP-1995
22 FILING DATE: 15-SEP-1995
23 FILING DATE: 15-SEP-1995
24 FILING DATE: 15-SEP-1995
25 FILING DATE: 15-SEP-1995
26 FILING DATE: 15-SEP-1995
27 FILING DATE: 15-SEP-1995
28 FILING DATE: 15-SEP-1995
29 FILING DATE: 15-SEP-1995
30 FILING DATE: 15-SEP-1995
31 FILING DATE: 15-SEP-1995
32 FILING DATE: 15-SEP-1995
33 FILING DATE: 15-SEP-1995
34 FILING DATE: 15-SEP-1995
35 FILING DATE: 15-SEP-1995
36 FILING DATE: 15-SEP-1995
37 FILING DATE: 15-SEP-1995
38 FILING DATE: 15-SEP-1995
39 FILING DATE: 15-SEP-1995
40 FILING DATE: 15-SEP-1995
41 FILING DATE: 15-SEP-1995
42 FILING DATE: 15-SEP-1995
43 FILING DATE: 15-SEP-1995
44 FILING DATE: 15-SEP-1995
45 FILING DATE: 15-SEP-1995
46 FILING DATE: 15-SEP-1995
47 FILING DATE: 15-SEP-1995
48 FILING DATE: 15-SEP-1995
49 FILING DATE: 15-SEP-1995
50 FILING DATE: 15-SEP-1995
51 FILING DATE: 15-SEP-1995
52 FILING DATE: 15-SEP-1995
53 FILING DATE: 15-SEP-1995
54 FILING DATE: 15-SEP-1995
55 FILING DATE: 15-SEP-1995
56 FILING DATE: 15-SEP-1995
57 FILING DATE: 15-SEP-1995
58 FILING DATE: 15-SEP-1995
59 FILING DATE: 15-SEP-1995
60 FILING DATE: 15-SEP-1995
61 FILING DATE: 15-SEP-1995
62 FILING DATE: 15-SEP-1995
63 FILING DATE: 15-SEP-1995
64 FILING DATE: 15-SEP-1995
65 FILING DATE: 15-SEP-1995
66 FILING DATE: 15-SEP-1995
67 FILING DATE: 15-SEP-1995
68 FILING DATE: 15-SEP-1995
69 FILING DATE: 15-SEP-1995
70 FILING DATE: 15-SEP-1995
71 FILING DATE: 15-SEP-1995
72 FILING DATE: 15-SEP-1995
73 FILING DATE: 15-SEP-1995
74 FILING DATE: 15-SEP-1995
75 FILING DATE: 15-SEP-1995
76 FILING DATE: 15-SEP-1995
77 FILING DATE: 15-SEP-1995
78 FILING DATE: 15-SEP-1995
79 FILING DATE: 15-SEP-1995
80 FILING DATE: 15-SEP-1995
81 FILING DATE: 15-SEP-1995
82 FILING DATE: 15-SEP-1995
83 FILING DATE: 15-SEP-1995
84 FILING DATE: 15-SEP-1995
85 FILING DATE: 15-SEP-1995
86 FILING DATE: 15-SEP-1995
87 FILING DATE: 15-SEP-1995
88 FILING DATE: 15-SEP-1995
89 FILING DATE: 15-SEP-1995
90 FILING DATE: 15-SEP-1995
91 FILING DATE: 15-SEP-1995
92 FILING DATE: 15-SEP-1995
93 FILING DATE: 15-SEP-1995
94 FILING DATE: 15-SEP-1995
95 FILING DATE: 15-SEP-1995
96 FILING DATE: 15-SEP-1995
97 FILING DATE: 15-SEP-1995
98 FILING DATE: 15-SEP-1995
99 FILING DATE: 15-SEP-1995
100 FILING DATE: 15-SEP-1995

```

```

PCT-US94-10261A-24
1  RESULT 6
2  PCT-US94-10261A-24
3  Sequence 24, Application PC/TUS9410261A
4  GENERAL INFORMATION:
5  APPLICANT: Carter, William G.
6  APPLICANT: Gil, Susanna C.
7  APPLICANT: Ryan, Maureen A.
8  TITLE OF INVENTION: Epiligrin, an Epithelial Ligand for
9  TITLE OF INVENTION: Integrins
10 NUMBER OF SEQUENCES: 30
11 CURRENT APPLICATION DATA:
12 FILING DATE: 02-SEP-1994
13 FILING DATE: 02-SEP-1994
14 FILING DATE: 02-SEP-1994
15 FILING DATE: 02-SEP-1994
16 FILING DATE: 02-SEP-1994
17 FILING DATE: 02-SEP-1994
18 FILING DATE: 02-SEP-1994
19 FILING DATE: 02-SEP-1994
20 FILING DATE: 02-SEP-1994
21 FILING DATE: 02-SEP-1994
22 FILING DATE: 02-SEP-1994
23 FILING DATE: 02-SEP-1994
24 FILING DATE: 02-SEP-1994
25 FILING DATE: 02-SEP-1994
26 FILING DATE: 02-SEP-1994
27 FILING DATE: 02-SEP-1994
28 FILING DATE: 02-SEP-1994
29 FILING DATE: 02-SEP-1994
30 FILING DATE: 02-SEP-1994
31 FILING DATE: 02-SEP-1994
32 FILING DATE: 02-SEP-1994
33 FILING DATE: 02-SEP-1994
34 FILING DATE: 02-SEP-1994
35 FILING DATE: 02-SEP-1994
36 FILING DATE: 02-SEP-1994
37 FILING DATE: 02-SEP-1994
38 FILING DATE: 02-SEP-1994
39 FILING DATE: 02-SEP-1994
40 FILING DATE: 02-SEP-1994
41 FILING DATE: 02-SEP-1994
42 FILING DATE: 02-SEP-1994
43 FILING DATE: 02-SEP-1994
44 FILING DATE: 02-SEP-1994
45 FILING DATE: 02-SEP-1994
46 FILING DATE: 02-SEP-1994
47 FILING DATE: 02-SEP-1994
48 FILING DATE: 02-SEP-1994
49 FILING DATE: 02-SEP-1994
50 FILING DATE: 02-SEP-1994

```

Query Match 17.7%: Score 64.5; DB 4; Length 1713;  
Best Local Similarity 34.4%: Pred. No. 9.2;  
Matches 21; Conservative 12; Mismatches 21; Indels 7; Gaps 3;  
QY 9 VQUSARGCOTSDNAAEALN-ASSKEALQ-IIEKIPKYYFOFINLNSKIDRIFD 65  
DB 553 VDAATAYENLAIKAEDANRAASASALQTVIKEDLPK-----AKTLSSNDKLLN 608  
QY 66 K 66  
DB 609 E 609  
RESULT 7  
US-07-745-206A-2  
Sequence 2, Application US/07745206A  
Patent No. 5429921  
GENERAL INFORMATION:  
APPLICANT: Harpold, Michael  
APPLICANT: Ellis, Steven  
APPLICANT: Williams, Mark  
APPLICANT: Feldman, Daniel  
APPLICANT: McCue, Ann  
TITLE OF INVENTION: Human Calcium Channel Compositions and Methods  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fitch, Even, Tabin & Flannery  
STREET: 135 S. LaSalle  
CITY: Chicago  
STATE: Illinois  
COUNTRY: U.S.A.  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/745,206A  
FILING DATE: 19910815  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitch, Even, Tabin & Flannery  
REFERENCE/DOCKET NUMBER: 51504  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-372-7842  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2161 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLCOLOGY TYPE: protein  
US-07-745-206A-2  
Query Match 17.4%: Score 63.5; DB 1; Length 2161;  
Best Local Similarity 28.4%: Pred. No. 17;  
Matches 21; Conservative 5; Mismatches 19; Indels 29; Gaps 1;  
QY 4 WGEFVDGSRGAGCGVYS-----NDNAEALNASK 34  
DB 707 WNAWVDGAGCGSSKVICVIFITLCGNYILLNVLFLAVLNLAJNESLNTAK 766  
QY 35 EEAQGLIIRKPEK 48

Query Match 17.4%; Score 63.5; DB 1; Length 2161;  
Best Local Similarity 28.4%; Pred. No. 17;  
Matches 21; Conservative 5; Mismatches 19; Indels 29; Gaps 1;

RESULT 8  
US-08-455-543A-49



REFERENCE/DOCKET NUMBER: 52516 (P519739)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619)238-0999  
INFORMATION FOR SEQUENCE 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2161 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: Internal  
US-08-223-305C-49

Query Match 17.4%; Score 63.5; DB 2; Length 2161;  
Best Local Similarity 28.4%; Pred. No. 17;  
Matches 21; Conservative 5; Mismatches 19; Indels 29; Gaps 1;

OY 4 WGEFVQGRSANGCQOTS-----NDAAEALNASK 34  
Db 707 NNAVYDGMAYGPPSSGMIVCIYFILFCGNYILLNVLAVNDLADRES/INTAK 766

OY 35 EEAQLIREKPEK 48  
Db 767 EEAERKKKIARK 780

RESULT 11  
US-08-223-305C-51  
Sequence 51, Application US/08223305C  
Patent No. 5851824  
GENERAL INFORMATION:  
APPLICANT: Harpold, Michael  
APPLICANT: Ellis, Steven  
APPLICANT: Williams, Mark  
APPLICANT: Williams, Daniel  
APPLICANT: McQue, Ann  
APPLICANT: Brenner, Robert  
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND  
NUMBER OF SEQUENCES: 57  
CORRESPONDENCE ADDRESS:  
CORRESPONDENCE ADDRESS: Harpold, Michael, Haller & McClain  
STREET: 1660 Union Street  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-2926

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSO Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/223.305C  
FILING DATE: April 4, 1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/469.354  
FILING DATE: Nov 1, 1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/745.206  
FILING DATE: 15-AUG-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/620.250  
FILING DATE: 30-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/482.384  
FILING DATE: 20-FEB-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/603.751  
FILING DATE: 04-APR-1989  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US89/01408  
FILING DATE: 04-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/176.899  
FILING DATE: 04-APR-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Seidman, Stephanie L.  
REGISTRATION NUMBER: 33,779  
REFERENCE/DOCKET NUMBER: 52516 (P519739)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619)238-0999  
TELEFAX: (619)238-0062  
INFORMATION FOR SEQUENCE 51:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2161 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: Internal  
US-08-223-305C-51

Query Match 17.4%; Score 63.5; DB 2; Length 2161;  
Best Local Similarity 28.4%; Pred. No. 17;  
Matches 21; Conservative 5; Mismatches 19; Indels 29; Gaps 1;

OY 4 WGEFVQGRSANGCQOTS-----NDAAEALNASK 34  
Db 707 NNAVYDGMAYGPPSSGMIVCIYFILFCGNYILLNVLAVNDLADRES/INTAK 766

OY 35 EEAQLIREKPEK 48  
Db 767 EEAERKKKIARK 780

RESULT 12  
US-08-311-363-2  
Sequence 2, Application US/08311363  
Patent No. 5875958  
GENERAL INFORMATION:  
APPLICANT: Harpold, Michael  
APPLICANT: Ellis, Steven  
APPLICANT: Williams, Mark  
APPLICANT: Williams, Daniel  
APPLICANT: McQue, Ann  
APPLICANT: Brenner, Robert  
TITLE OF INVENTION: Human Calcium Channel Compositions and  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
CORRESPONDENCE ADDRESS: Harpold, Michael, Haller & McClain  
STREET: 1660 Union Street  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-2926  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/311.363  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/745.206  
FILING DATE: 15-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Seidman, Stephanie L.  
REGISTRATION NUMBER: 33,779  
REFERENCE/DOCKET NUMBER: 6362-51506  
TELECOMMUNICATION INFORMATION:

	Query Match	15.9%	Score 58	DB 3	Length 844
	Best Local Similarity	27.5%	Pred. No. 26		
	Matches 11	Conservative 13	Mismatches 16	Indels	
QY	24	AAALNAAASKEAFQIIRKIPKYLFOFHNNSNDIR	63		
DB	714	ANHTTSLVATVEQALEMSNIIQIETENASSYVDVO	753		
IDENT	15				

```

US-07-637-870-9
: Sequence 9, Application US/07637870
: Patent No. 5310665
: GENETIC INFORMATION
: APPLICANT: Lasters, Anne-Marie
: APPLICANT: Lasters, Ignace
: APPLICANT: Quax, Wilhemus J.
: APPLICANT: Van Der Laan, Jan M.
: TITLE OF INVENTION: NOVEL GLUCOSE ISOMERASES HAVING
: TITLE OF INVENTION: ALTERED SUBSTRATE SPECIFICITY
: NUMBER OF SEQUENCES: 9
: COMPLETION ADDRESS:
: ADDRESSEE: Middelburgse Poort
: STREET: 545 Middlefield Road, Suite 200
: CITY: Menlo Park
: STATE: California
: COUNTRY: USA
: ZIP: 94025
: COMPUTER READABLE FORM:
: MODULATED: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/637-870
: FILING DATE: 19910104
: CLASSIFICATION: C12N 33
: ATTORNEY/AGENT INFORMATION:
: NAME: Murrishioe, Kate H.
: REGISTRATION NUMBER: 29,959
: REFERENCE/DOCKET NUMBER: 24615-20019.00
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-813-5600
: TELEFAX: 415-327-2951
: ALEXANDER PAPER CO. ID NO: 9:
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 439 amino acids
: TYPE: AMINO ACID
: STRANDEDNESS: single
: TOPOLOGY: linear
US-07-637-870-9

Query Match 15.88; Score 57.5; DB 1; Length 439;
Best Local Similarity 25.68; Pred. No. 13;
Matches 20; Conservative 11; Mismatches 30; Indels 17; Gaps 3;

QY 4 WGEFVQDGRARGCOTSN-----DNAAEALNASKEEALQIIRKIPKYLQFHNLN-- 57
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 48 WHITMGSDPPGATNGFNPWDEUGSELDNRHRVNFETAEKLGKVI-ICFHDIDIA 106
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 58 -----SNLDRIFD 65
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 107 PTGNSLKRFYANLDEITD 124
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Search completed: February 3, 2001, 02:19:38
Job time: 1700 sec

```



GenCore version 4.5  
 Copyright (c) 1993 - 2000 CompuGen Ltd.  
 OM protein - protein search, using sw model  
 Run on: February 3, 2001, 01:18:13 ; Search time 144.12 Seconds  
 (without alignments)  
 16,608 Million cell updates/sec

Title: US-09-289-346a-1  
 Perfect score: 364  
 Sequence: 1 TLWGVEPOVDSRGSGCOT.....PQEHNLNLDIRFDKTEP 70

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_36.\*

1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
 2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
 3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
 4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
 5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
 6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
 7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
 8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
 9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
 10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
 11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
 12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
 13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
 14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
 15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
 16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
 17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
 18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
 19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
 20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
 21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	230	63.2	353	18 W34338	Bean golden mosaic
2	230	63.2	353	18 W34332	Bean golden mosaic
3	230	63.2	353	18 W34333	Bean golden mosaic
4	230	63.2	353	18 W34334	Bean golden mosaic
5	230	63.2	353	18 W34335	Bean golden mosaic
6	230	63.2	353	17 W34336	Sardinian tomato y
7	230	63.2	359	17 P88871	Sardinian tomato y
8	230	63.2	359	17 P88872	Sardinian tomato y
9	228	62.6	353	8 P70407	ORF 4 gene product
10	228	62.6	361	18 W34336	Tomato mottle virus
11	228	62.6	361	18 W34334	Tomato mottle virus
12	228	62.6	361	18 W34325	Tomato mottle virus

13	228	62.6	361	18 W34326	Tomato mottle virus
14	221.5	60.9	361	8 P70562	Product of ORF 4 f
15	216	59.3	362	19 W56495	Tobacco leaf curl
16	215	59.1	357	18 W34329	Tomato yellow leaf
17	213	58.1	357	18 W34330	Tomato yellow leaf
18	211	59.1	357	18 W34331	Tomato yellow leaf
19	207	56.9	357	18 W34337	Tomato yellow leaf
20	67.5	18.5	512	19 W68473	HIV-1 strain YBF30
21	65	17.9	292	21 Y92317	Geminivirus Rep C1
22	65	17.9	335	21 Y92318	Mastrevirus Rep in
23	64.5	17.7	299	18 W54540	Reduced sequence o
24	64.3	17.4	1719	15 P20146	Sequence of the al
25	63.5	17.4	1719	15 P20146	Sequence of the al
26	63.5	17.4	2161	14 K33555	Human neuronal cal
27	63.5	17.4	2161	16 R71001	Human neuronal cal
28	63.5	17.4	2161	16 R71002	Human neuronal cal
29	63.5	17.4	2161	19 W63137	Human calcium chan
30	63.5	17.4	2161	19 W63137	Human calcium chan
31	62.5	17.2	447	21 P87089	Human secreted pro
32	62.5	17.2	447	21 P87089	Membrane-bound pro
33	62.3	17.2	447	21 Y50947	Human secreted pro
34	62.5	17.2	456	21 Y87190	Human secreted pro
35	60.5	16.6	131	18 W34320	Tomato mottle virus
36	60.5	16.6	204	20 W89814	Protein encoded by
37	60.5	16.6	299	20 W89823	Protein encoded by
38	60.5	16.6	299	20 W89849	Antigen from clust
39	58.5	16.1	297	8 P70318	Protein M1 encoded
40	58.5	16.1	299	19 W93322	H. pylori GPHO 136
41	58.5	16.1	299	19 W93322	H. pylori GPHO 136
42	58	15.9	318	20 Y31895	Sorbeta cyclin del
43	58	15.9	318	20 Y31895	Sorbeta cyclin del
44	57.5	15.8	844	18 W03773	Histidinol dehydro
45	57	15.7	108	16 Y95043	Candida albicans p
					Murine mutant cyto

## ALIGNMENTS

RESULT 1  
 ID W34338 standard; Protein: 353 AA.  
 XX  
 XX W34338:  
 XX  
 XX 27-APR-1998 (first entry)  
 XX DE Bean golden mosaic geminivirus C1 protein.  
 XX DE Geminivirus: BCM; C1 gene: transdominant mutation;  
 KW transgenic plant; disease resistance.  
 KW  
 OS Bean golden mosaic virus type II isolate Guatemala.  
 XX  
 PN W09739110-A1.  
 XX  
 PD 23-OCT-1997.  
 XX  
 XX 15-APR-1997: 97WO-US05300.  
 XX  
 XX 16-APR-1996: 96US-0015517.  
 XX  
 XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.  
 XX (WISC ) WISCONSIN ALUMNI RES FOUND.  
 XX  
 XX Abiquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;  
 XX WPI: 1997-526447/48.  
 XX N-FSDB, P93314.  
 XX  
 XX Transgenic plants expressing geminivirus AC1 and CI wild-type and  
 PT mutant genes - have increased resistance to geminivirus infection  
 PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean  
 PT golden mosaic geminivirus

XX Example 5; Page 100-102; 132pp; English.

XX This sequence comprises the bean golden mosaic virus (BGMV) C1  
XX protein that is required for replication. The invention involves  
XX production of transgenic plants containing DNA comprising C1 or AC1  
XX wild-type or mutant sequences that negatively interfere in trans  
XX with geminiviral replication during infection. Such transgenic  
XX plants are resistant to viral infection. The AC1/C1 genes are  
XX especially from BGMV, tomato mottle virus or tomato yellow leaf  
XX curl virus (see T93282-93) and encode polypeptides (see W34324-35)  
XX and/or the NTP-binding domains in the highly conserved DNA-nicking domain  
XX and/or the NTP-binding domains.

XX Sequence 353 AA;

Query Match 63.2%; Score 230; DB 18; Length 353;

Best Local Similarity 61.4%; Pred. No. 6, 5c-21;

Matches 43; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

QY 1 TLVNGFEQVDRSGRGCGTSDNDAAEALNASSKEALQIREKIPKYLFOFINLNSNL 60

Db 110 TLVNGFEQVDRSGRGCGTSDNDAAEALNASSKEALQIREKIPKYLFOFINLNSNL 60

QY 61 DRIFOKTPEP 70

Db 170 CRIFVKYPEP 179

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

CC wild-type or mutant sequences that negatively interfere in trans  
CC with geminiviral replication during infection. Such transgenic  
CC plants are resistant to viral infection. The AC1/C1 genes are  
CC especially from BGMV, tomato mottle virus or tomato yellow leaf  
CC curl virus (see T93282-93) and encode polypeptides (see W34324-35)  
CC and/or the NTP-binding domains in the highly conserved DNA-nicking and/or the  
CC NTP-binding domains.

XX Sequence 353 AA;

Query Match 63.2%; Score 230; DB 18; Length 353;

Best Local Similarity 61.4%; Pred. No. 6, 5c-21;

Matches 43; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

QY 1 TLVNGFEQVDRSGRGCGTSDNDAAEALNASSKEALQIREKIPKYLFOFINLNSNL 60

Db 110 TLVNGFEQVDRSGRGCGTSDNDAAEALNASSKEALQIREKIPKYLFOFINLNSNL 60

QY 61 DRIFOKTPEP 70

Db 170 CRIFVKYPEP 179

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111

XX 1111111111



OY 61 DRIFDKTPEP 70  
 Db 170 grlrvkvpap 179

RESULT 6  
 R88870 ID R88970 standard; Protein: 359 AA.

XX AC R88870;  
 XX DT 07-NOV-1996 (first entry)

XX Sardinian tomato yellow leaf curl virus mutated Rep protein (K227A).  
 XX Nucleotide triphosphate binding site; DNA helicase; RNA helicase;  
 KW modification; mutation; viral replication; deficient; inhibition;  
 KW viral resistance; geminivirus; tomato yellow leaf curl virus;  
 KW Sardinian isolate; STYLVCV; transgenic plant; P-loop; CI protein;  
 KW AI1 protein; dominant negative phenotype.  
 XX OS Sardinian tomato yellow leaf curl virus.

XX Key Location/Qualifiers  
 FT Misc-difference 227 /note= "wild-type Lys has been replaced by Ala"  
 XX PN W09608573-Al.

XX PD 21-MAR-1996.  
 XX PF 15-SEP-1995; 95WO-FR01192.  
 XX PL 15-SEP-1994; 94FR-0011040.  
 XX PA (CNRS ) CENT NAT RECH SCI.  
 XX PI Gronenborn B;

XX WPI; 1996-179947/18.  
 XX DR N-PSDB: T12904.  
 XX PT Prodn. of virus-resistant transgenic plants - using mutated genomic  
 PT sequence from phytopathogenic DNA virus  
 XX PS Disclosure; Fig 13; 93pp; French.

XX Mutation of consensus amino acids in the RFP-binding site of  
 CC geminivirus Rep protein is used to produce replication deficient  
 CC viruses. The mutated viral nucleic acid is used for producing  
 CC transgenic plants that are resistant to, or tolerant of, the native  
 CC virus. The present sequence is a mutant form of the Rep (or CI)  
 CC protein from the Sardinian isolate of tomato yellow leaf curl virus  
 CC (STYLVCV) in which the wild-type Lys227 residue has been changed to an  
 CC Ala residue; transgenic Nicotiana benthamiana plants generated by  
 CC transformation with the mutated virus were found to be resistant to  
 CC STYLVCV, i.e. the mutation results in a dominant negative phenotype.

XX Sequence 359 AA;

Query Match 63.2%; Score 230; DB 17; Length 359;  
 Best Local Similarity 59.4%; Pred. No. 6.6e-21;  
 Matches 41; Conservative 11; Mismatches 17; Indels 0; Gaps 0;

OY 2 LVWGEFVDGSRAGCGCOTSNDAAEALMASSEKALQIIREKIPKYLEFQHNLNSLD 61  
 Db 111 lvgwtfgldgrsar99gqtdandayakainagksqaldvikelaprdyvlhfhmnsld 170

OY 62 RIFDKTPEP 70

Db 171 kvfqvppap 179

RESULT 8

RESULT 7

R88870 ID R88871 standard; Protein: 359 AA.  
 XX AC R88871;  
 XX DT 07-NOV-1996 (first entry)

XX Sardinian tomato yellow leaf curl virus mutated Rep protein (K227H).  
 XX Nucleotide triphosphate binding site; DNA helicase; RNA helicase;  
 KW modification; mutation; viral replication; deficient; inhibition;  
 KW viral resistance; geminivirus; tomato yellow leaf curl virus;  
 KW Sardinian isolate; STYLVCV; transgenic plant; P-loop; CI protein;  
 KW AI1 protein.

XX OS Sardinian tomato yellow leaf curl virus.

XX Key Location/Qualifiers  
 FT Misc-difference 227 /note= "wild-type Lys has been replaced by His"

XX PN W09608573-Al.  
 XX PD 21-MAR-1996.  
 XX PF 15-SEP-1995; 95WO-FR01192.  
 XX PL 15-SEP-1994; 94FR-0011040.  
 XX PA (CNRS ) CENT NAT RECH SCI.

XX Gronenborn B;

XX WPI; 1996-179947/18.

XX DR N-PSDB: T12905.

XX PT Prodn. of virus-resistant transgenic plants - using mutated genomic  
 PT sequence from phytopathogenic DNA virus

XX PS Disclosure; Fig 13; 93pp; French.

XX Mutation of consensus amino acids in the RFP-binding site of  
 CC geminivirus Rep protein is used to produce replication deficient  
 CC viruses. The mutated viral nucleic acid is used for producing  
 CC transgenic plants that are resistant to, or tolerant of, the native  
 CC virus. The present sequence is a mutant form of the Rep (or CI)  
 CC protein from the Sardinian isolate of tomato yellow leaf curl virus  
 CC (STYLVCV) in which the wild-type Lys227 residue has been changed to a  
 CC Ala residue; transgenic Nicotiana benthamiana plants generated by  
 CC transformation with the mutated virus were not resistant to STYLVCV.  
 CC In contrast, plants transformed with a virus in which Lys227 had been  
 CC replaced by Ala were found to be resistant.

XX Sequence 359 AA;

Query Match 63.2%; Score 230; DB 17; Length 359;  
 Best Local Similarity 59.4%; Pred. No. 6.6e-21;  
 Matches 41; Conservative 11; Mismatches 17; Indels 0; Gaps 0;

OY 2 LVWGEFVDGSRAGCGCOTSNDAAEALMASSEKALQIIREKIPKYLEFQHNLNSLD 61  
 Db 111 lvgwtfgldgrsar99gqtdandayakainagksqaldvikelaprdyvlhfhmnsld 170

OY 62 RIFDKTPEP 70

Db 171 kvfqvppap 179

AC	P70407;
XX	
DT	02-MAY-1991 (first entry)
XX	
DE	ORF 4 gene product of Bean Golden Mosaic virus.
XX	
DE	plant vector.
XX	
KW	Bean golden mosaic virus.
XX	
OS	JP61257186-A.
XX	
PN	
XX	
PD	14-NOV-1986.
XX	
FF	10-MAY-1985; 85JP-0098108.
XX	
PR	10-MAY-1985; 85JP-0098108.
XX	
PA	(TEIJ) TEIJJN KK.
XX	
NFL	NFL: 1987-159662/23.
DR	N-P5DB; N70630.
XX	
XX	New DNA and hybrid DNA - used for recombinant vector of plants.
PS	Disclosure: Fig 6; 24pp: Japanese.
CC	The sequence encoding this protein may be taken along with the -a
CC	DNA sequence and a suitable resistance gene, and used to create a
CC	recombinant plant vector.
CC	See also N70629.
XX	
CC	
XX	
SQ	Sequence 353 AA;
Query Match	62.6%; Score 228; DB 8; Length 353;
Best Local Similarity	61.4%; Prod. No. 1.2e-20;
Matches 43	Conservative 13; Mismatches 14; Indels 0; Gaps
Qy	1 TLAKGEYDVRGSRAGCGDSNDAAAFALNASSKSEALOTIRPKIDKYVLFGPINLSWL 60
Db	110 :     :     :     :     :     :     :     :     :     :
	110 tiewgfgvdrsgarqggandayakalnadsalsitlikeepkdyvlghairsnl 169
Qy	61 DRIFDKTPEP 70
	:
Db	170 eriffkvpep 179
	:
RESULT 10	
ID	M34336 standard; Protein: 361 AA.
XX	
AC	M34336;
XX	
TT	Tomato mottle virus AC1 protein.
XX	
DE	27-APR-1998 (first entry)
XX	
KW	Geminivirus; TOMOV; AC1 gene; transdominant mutation;
KW	transgenic plant; disease resistance.
XX	
OS	Tomato mottle virus isolate Florida.
XX	
PN	W09739110-A1.
XX	
PD	23-OCT-1997.
XX	
PF	15-APR-1997; 97WO-US06300.
XX	
XX	
XX	16-APR-1996; 96US-001517.
XX	
PA	(SEM-) SEMINIS VEGETABLE SEEDS, INC.
PA	(WISC.) WISCONSIN AGRIMET SEEDS FOUND.

XX PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;  
 XX WP1: 1997-526447/48.  
 XX N-PSDB: T93282.

XX Transgenic plants expressing geminivirus AC1 and CI wild-type and  
 PT mutant genes - have increased resistance to geminivirus infection  
 PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean  
 PT golden mosaic geminivirus

XX Example 3.4: Page 57-58; 132pp; English.

XX This protein comprises the wild-type AC1 protein of tomato mottle  
 CC virus (ToMoV), a geminivirus that has a bipartite genome. The AC1  
 CC gene (see T93294), must be expressed for efficient replication of  
 CC the two genomic components, DNA-A and DNA-B. The AC1 protein has a  
 CC DNA binding site specific to the DNA-A common region, a DNA nicking  
 CC activity, and an NTP binding activity. The invention involves  
 CC production of transgenic plants containing DNA comprising AC1 or CI  
 CC genes that are resistant to viral infection. Such transgenic  
 CC plants are resistant to viral infection. The AC1/CI genes are  
 CC especially from ToMoV, tomato yellow leaf curl virus or bean golden  
 CC mosaic geminivirus (see T93282-93) and encode polypeptides (see  
 CC W34324-35) that have mutations in the highly conserved DNA-nicking  
 CC domain and/or the NTP-binding domain.

XX Sequence 361 AA;

Query Match 62.6%; Score 228; DB 18; Length 361;  
 Best Local Similarity 58.6%; Pred. No. 1.2e-20;  
 Matches 41; Conservative 16; Mismatches 13; Indels 0; Gaps 0;

QY 1 TLVNGEFQVDRSGRGQTSNDAAALNASSKEALQITREKTEPKYLFQFHNLSNL 60  
 DB 110 LIEWGDFQIGRSRGQTSNDAAALNASSKEALQITREKTEPKYLFQFHNLSNL 169  
 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 QY 61 DRIFQKTPPEP 70  
 DB 170 erifakapep 179  
 :|||:|||||

RESULT 11  
 W34324  
 ID W34324 standard; Protein; 361 AA.  
 XX W34324;  
 AC W34324;  
 DT 27-APR-1998 (first entry)

XX Tomato mottle virus AC1 mutant ToMoV-AC1dml.  
 DE Geminivirus: ToMoV-AC1dml: AC1 gene; transdominant mutation:  
 KW transgenic plant; disease resistance.

XX Tomato mottle virus isolate Florida.  
 OS Synthetic.

XX W09739110-A1.

PD 23-OCT-1997.

XX 15-APR-1997; 97MO-US06300.

PR 16-APR-1996; 96US-0015517.

XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.  
 PA (WISC) WISCONSIN ALLUMNI RES FOUND.

PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;

XX

DR WP1: 1997-526447/48.  
 DR N-PSDB: T93282.

XX Transgenic plants expressing geminivirus AC1 and CI wild-type and  
 PT mutant genes - have increased resistance to geminivirus infection  
 PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean  
 PT golden mosaic geminivirus

XX Example 3.4: Page 60-62; 132pp; English.

XX This protein comprises a transdominant lethal mutant, designated  
 CC ToMoV-AC1dml, of tomato mottle virus (ToMoV) AC1 protein (see  
 CC W34324) and a transdominant lethal mutant AC1 gene (see  
 CC W34325) that are co-expressed by ToMoV-AC1dml. The AC1 gene in its  
 CC NTP-binding domains. The AC1 gene (see also T93294) must be  
 CC expressed for efficient replication of the two genomic components,  
 CC DNA-A and DNA-B, of the bipartite ToMoV genome. The invention  
 CC involves production of transgenic plants containing DNA comprising  
 CC geminivirus AC1 or CI wild-type or mutant sequences that negatively  
 CC interfere in trans with geminiviral replication during infection.  
 CC Such transgenic plants are resistant to viral infection. Such  
 CC AC1/CI genes are especially from ToMoV, tomato yellow leaf curl  
 CC virus or bean golden mosaic geminivirus (see T93282-93) and encode  
 CC polypeptides (see W34324-35) that have mutations in the highly  
 CC conserved DNA-nicking and/or NTP-binding domains.

XX Sequence 361 AA;

Query Match 62.6%; Score 228; DB 18; Length 361;  
 Best Local Similarity 58.6%; Pred. No. 1.2e-20;  
 Matches 41; Conservative 16; Mismatches 13; Indels 0; Gaps 0;

QY 1 TLVNGEFQVDRSGRGQTSNDAAALNASSKEALQITREKTEPKYLFQFHNLSNL 60  
 DB 110 LIEWGDFQIGRSRGQTSNDAAALNASSKEALQITREKTEPKYLFQFHNLSNL 169  
 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 QY 61 DRIFQKTPPEP 70  
 DB 170 erifakapep 179  
 :|||:|||||

RESULT 12  
 W34325  
 ID W34325 standard; Protein; 361 AA.  
 XX W34325;  
 AC W34325;  
 DT 27-APR-1998 (first entry)

XX Tomato mottle virus AC1 mutant ToMoV-AC1dml.

XX Geminivirus: ToMoV-AC1dml: AC1 gene; transdominant mutation:  
 KW transgenic plant; disease resistance.

XX Tomato mottle virus isolate Florida.  
 OS Synthetic.

XX W09739110-A1.

PD 23-OCT-1997.

XX 15-APR-1997; 97MO-US06300.

PR 16-APR-1996; 96US-0015517.

XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.  
 PA (WISC) WISCONSIN ALLUMNI RES FOUND.

PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;

XX WP1: 1997-526447/48.  
 DR N-PSDB: T93282.

mutant genes – have increased resistance to geminivirus infection  
e.g. tomato mottle virus, tomato yellow leaf curl virus or bean  
golden mosaic geminivirus

example 3.6: Page 67-69; 12pp; English.

This protein comprises a transdominant lethal mutant, designated  
Tomov-Ac1ind23, of tomato mottle virus (Tomov) Ac1 protein (see  
W3436). It is encoded by a mutant Ac1 gene (see T93284) of  
tomato mottle virus. The Ac1 gene carries 2 mutations in an  
NTP-binding domain. The Ac1 gene (see also T93242) must be  
expressed for efficient replication of the two genomic components,  
DNA-A and DNA-B, of the bipartite Tomov genome. The invention  
relates to the use of the mutant Ac1 protein for controlling  
geminivirus Ac1 or C1 wild-type or mutant sequences that negatively  
interfere in trans with geminiviral replication during infection.  
Such transgenic plants are resistant to viral infection. The  
Ac1/AC1 genes are especially from Tomov, tomato yellow leaf curl  
virus or bean golden mosaic geminivirus (see T93282-93) and encode  
polypeptides (see W34324-35) that have mutations in the highly  
conserved DNA-binding and/or NTP-binding domains.

Sequence 361 AA:

```

Query Match      62.6%  Score 228:  DB 18:  Length 361:
Best Local Similarity 58.6%  Pred. No. 1,2e-20:
Matches 41:  Conservative 16:  Mismatches 13:  Indels 0:  Gaps
QY  1  TAWGEEQVDSRARGCQTSNDMAALNASSKKEALQIREKIPKYLQFHNNLNL 60
Db   | 11111111111111111111111111111111111111111111111111111
110  |11111111111111111111111111111111111111111111111111111
      |11111111111111111111111111111111111111111111111111111
QY  61  DRIFQKPTKPEP 70
Db   | 11111111111111111111111111111111111111111111111111111
QY  170  erifakapep 179
      | 11111111111111111111111111111111111111111111111111111

```

PD	7P0562 standard; Protein; 361 AA.
PC	XX
AC	P70552;
CC	XX
XX	DT 30-APR-1991 (first entry)
XX	DE Product of ORF 4 from MTMV complementary strand (b).
XX	XX
XX	XX Geminivirus.
XX	XX
XX	XX Mungbean yellow mosaic virus.
XX	XX
PN	JP62136982-A.
XX	XX
XX	XX 09-JUN-1987.
PD	XX
XX	XX
PF	XX 28-NOV-1985; 85JP-0266080.
XX	XX
PR	XX 28-NOV-1985; 85JP-0266080.
XX	XX
XX	XX (TEIJ ) TEIJIN KK.
XX	XX
XX	XX WPI: 1987-196308/28.
DR	XX
DR	XX N-PSDB: N70895.
PT	XX Novel DNA and hybrid DNA useful - as vector for recom-
PT	XX of plant gene..
XX	XX
PS	XX Disclosure; Fig 8, 2lpp; Japanese.
XX	XX
XX	XX The sequence is encoded by ORF 4, which occurs on the
XX	XX strand of the (b) molecule of the geminivirus.
CC	XX See also: P70559-P70566.
CC	XX

PT Novel DNA and hybrid DNA useful - as vector for recombinant work  
PT of plant gene.  
XX  
XX  
XX Disclosure; Fig 8; 21pp; Japanese.  
XX  
XX The sequence is encoded by ORF 4 which occurs on the complementary  
CC strand of the (b) molecule of the geminivirus.  
CC See also. p70549-p70567

```

XX SQ Sequence 361 AA;
      Query Match 60.9%; Score 221.5; DB 8; Length 361;
      Best Local Similarity 67.7%; Pred No. 7, 8e-20;
      Matches 44; Conservative 9; Mismatches 11; Indels 1; Gaps 1;

QY 5 GEFQVDSRGSGCQTSNDAAAEALNASKEEALQIREKIPEKYLFOFHNLNSLDRIE 64
   | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 114 gsfqvdgrsrgsgkpsandayaaealngsklqalnllrekakpkeyllqfhlncnlsrlf 173
QY 65 -DKTP 68
   | : |
Db 174 adavp 178

RESULT 15
W56495 standard; Protein; 362 AA.
XX W56495;
AC W56495;
DT 11-AUG-1998 (first entry)
DE Tobacco leaf curling virus gene protein Cl.
KW Tobacco leaf curling virus gene; TLCV; promoter; Cl protein.
OS Tobacco leaf curling virus.
XX JP10070982-A.
XX 17-MAR-1998.
PF 30-AUG-1996; 96JP-0230394.
PR 30-AUG-1996; 96JP-0230394.
XX (NORQ ) NORINSUTSANSO KYUSHU NOGYO SHIKENJO.
PA WPI; 1998-233630/21.
XX Tobacco leaf curling virus gene - useful for inserting into vectors
PT for expression in, e.g. tomato plants
XX Claim 5; Fig 5; 9pp; Japanese.
XX This sequence represents the Cl protein encoded by the tobacco leaf
CC the TLCV gene of the invention. The gene or its promoter can
CC be inserted into vectors for expression in plants, e.g. tobacco
CC tomato. This sequence is believed to be encoded by the TLCV gene shown in
CC V29761.
XX SQ Sequence 362 AA;
      Query Match 50.3%; Score 216; DB 19; Length 362;
      Best Local Similarity 53.6%; Pred No. 3, 8e-19;
      Matches 45; Conservative 10; Mismatches 13; Indels 16; Gaps 1;

QY 2 LVNGEFQVDSRGSGCQTSNDAAAEALNASKEEALQIREKIPEKYLFOFHNLNSLD 61
   | : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111 ldfgfvqvdgrsarcgccsandayaaealngskssalnllrekakpfdvlfqhlnsld 170
QY 62 RI-----FKTPE 69
   | : ||
Db 171 rifaplevfcvfsssfddgpe 194

```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.  
OM protein - protein search, using sw model  
Run on: February 3, 2001, 02:22:50 ; Search time 180.33 Seconds  
(without alignments)  
45.497 Million cell updates/sec

Title: US-09-289-346a-2  
Perfect score: 362  
Sequence: 1 TLVWGFQVQGAAGAGCOT.....FQPHNLSNLDIFDKTPEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117027915 residues  
Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL15.\*  
1: sp.archaea.\*  
2: sp.bacteria.\*  
3: sp.fungi.\*  
4: sp.human.\*  
5: sp.invertebrate.\*  
6: sp.manual.\*  
7: sp.mammal.\*  
8: sp.organella.\*  
9: sp.phage.\*  
10: sp.plant.\*  
11: sp.todent.\*  
12: sp.virus.\*  
13: sp.vertic.\*  
14: sp.unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	281	77.6	226	12 009727	009727 leonurus mo
2	281	77.6	226	12 09WHF6	09WHF6 tomato mild
3	279	77.1	361	12 067574	067574 bean golden
4	277	76.5	225	12 09QDB1	09QDB1 cowpea gold
5	274	75.7	195	12 09B693	09B693 sugar potato
6	274	75.7	149	12 09B975	09B975 macropitiliu
7	266	73.5	233	12 09Y1A4	09Y1A4 macropitiliu
8	266	73.5	233	12 09Y1A4	09Y1A4 potato yell
9	260	71.8	234	12 093180	093180 tobacco lea
10	257	71.0	190	12 092084	092084 tobacco lea
11	257	71.0	190	12 092084	092084 tobacco lea
12	256	70.7	190	12 09W827	09W827 tobacco lea
13	256	70.7	208	12 0920A0	0920A0 tobacco lea
14	256	70.2	208	12 0920B8	0920B8 tobacco lea
15	254	70.2	208	12 0920B8	0920B8 tobacco lea
16	252	69.5	208	12 0920B6	0920B6 tobacco lea
17	251	69.3	203	12 092083	092083 tobacco lea
18	251	69.3	363	12 073577	073577 cotton leaf
19	244	67.4	208	12 0920A0	0920A0 tobacco lea

## SUMMARIES

```

20 244 67.4 363 12 072705
21 244 67.4 363 12 072719
22 243 67.1 190 12 0920B6
23 243 67.1 208 12 0920C6
24 242 66.9 362 12 0686B6
25 242 66.9 362 12 0686B6
26 241 66.6 359 12 092V24
27 241 66.6 359 12 092V24
28 241 66.6 359 12 092V27
29 241 66.6 359 12 092V27
30 240 66.3 353 12 072692
31 240 66.3 353 12 0686B8
32 239 66.0 190 12 0920A7
33 239 66.0 190 12 0920A7
34 239 66.0 362 12 091N48
35 238 65.7 307 12 091E77
36 238 65.7 361 12 072723
37 237 65.5 231 12 095620
38 237 65.5 354 12 091N42
39 236 65.2 358 12 065418
40 236 65.2 358 12 065418
41 232 64.1 362 12 092V27
42 232 64.1 363 12 073494
43 230 63.5 306 12 091E71
44 230 63.5 358 12 09WR17
45 230 63.5 358 12 09JEA9
072705 cotton leaf
072719 cotton leaf
0920B6 tobacco lea
0920C6 tobacco lea
0686B6 tomato paeu
0686B6 tomato yell
092V24 tomato yell
092V27 tomato yell
092V27 tomato yell
072692 beet curly
0686B8 beet curly
0686B8 beet curly
0920A7 tobacco lea
0920A7 tobacco lea
091N48 okra enatio
091E77 cotton leaf
072723 cotton leaf
095620 african tom
091N42 south afric
065418 okra rose
065418 okra rose
092V27 althea rose
073494 okra yellow
091E71 cotton leaf
09WR17 african cas
09JEA9 cassava gem

```

## ALIGNMENTS

```

RESULT 1
009727 PRELIMINARY: PRT: 226 AA.
ID 009727
AC 01-JUL-1997 (TRENDELrel. 04, Created)
DC 01-JUL-1997 (TRENDELrel. 04, Last sequence update)
DT 01-OCT-2000 (TRENDELrel. 16, Last annotation update)
DE REP PROTEIN (FRAGMENT).
GN REP.
OS Leonurus mosaic virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=58177;
PC (1)
RC SEQUENCE FROM N.A.
RA Farla J.C., Maxwell D.P.:
RL Submitted (Mar-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U92532; AA051157.1;
DR INTERPRO: IPR001191;
DR INTERPRO: IPR001301;
DR PROSITE: PS00227;
DR PRINTS: PR00227; GEMCOTALL1.
DR PRINTS: PR00228; GEMCOTALLV1.
FT NON_TER 226 226
SQ SEQUENCE 226 AA: 25617 MW: 73CDB6E7650B3FC5 CRC64:

```

Query Match 77.6%; Score 281; DB 12; Length 226;  
Local Similarity: 77.1%; P: 6.7e-43;  
Matches 54; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

```

Oy 1 TLVWGFQVQGAAGAGCOTSNDAAEALNASSKEALQIIRKIPKYLQFPHNLSNI 60
Db 11 TLVWGFQVQDSRRSGQQTVDNAAEALNAPDKRTALQIIRKIPKYLQFPHNLSNI 170
Oy 1 TLVWGFQVQDAAGAGCOTSNDAAEALNASSKEALQIIRKIPKYLQFPHNLSNI 60
Db 11 TLVWGFQVQDSRRSGQQTVDNAAEALNAPDKRTALQIIRKIPKYLQFPHNLSNI 170
Oy 61 DRIFDKTPEP 70
Db 171 DRIFAKAPEP 180

```

RESULT 2  
09WHF6



RA Lohrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.:  
 RT Detection of a geminivirus infecting sweet potato in the United  
 States.  
 RI PLoS ONE 2: 82-1257 (1998).  
 DR EMBL: AF104036; AAD47173.1; -  
 DR INTERPRO: IPR001191; -  
 DR INTERPRO: IPR001191; -  
 DR PFAM: PF00799; Gemin1\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATL1.  
 SQ SEQUENCE 364 AA; 40680 MW; 5F79752431A09D6E CRC64;

Query Match 76.5% Score 277; DB 12; Length 364;  
 Best Local Similarity 79.4% Pred. No. 3.1e-22;  
 Matches 54; Conservative 4; Mismatches 10; Indels 0; Gaps 0;  
 Qy 1 TUVGFEVQGAAGCGCOTNDAAEAALNASKEEALQIREKIPKYLQFPHNLSNU. 60  
 Db 110 TIVGFEVQDGSANGCGQTNDAAEAALNASKEEALQIREKIPKYLQFPHNLSNU. 169  
 Qy 61 DRFDKTP 68  
 Db 170 DRFSPPP 177

RESULT 6  
 Q98693 PRELIMINARY: PRT; 185 AA.  
 ID Q98693;  
 AC Q98693;  
 DT 01-FEB-1997 (TRENBLrel. 02, Created)  
 DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)  
 DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)  
 DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).  
 OS Macrotellium golden mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=51034;  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-JAMAICA;  
 RA Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;  
 RI Molecular characterization of two distinct geminiviruses infecting M.  
 alalytoides in Jamaica. (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U67526; AAB57865.1; -  
 DR INTERPRO: IPR001191; -  
 DR INTERPRO: IPR001191; -  
 DR PFAM: PF00799; Gemin1\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATL1.  
 FT NON\_TER 185 185  
 SQ SEQUENCE 185 AA; 20975 MW; 39113850A025A5EE1 CRC64;

Query Match 75.7% Score 274; DB 12; Length 185;  
 Best Local Similarity 70.0% Pred. No. 3e-22;  
 Matches 49; Conservative 13; Mismatches 8; Indels 0; Gaps 0;  
 Qy 1 TUVGFEVQGAAGCGCOTNDAAEAALNASKEEALQIREKIPKYLQFPHNLSNU. 60  
 Db 89 TIVGFEVQDGSANGCGQTNDAAEAALNASKEEALQIREKIPKYLQFPHNLSNU. 148  
 Qy 61 DRFDKTP 70  
 Db 149 DRFSKPP 158  
 RESULT 7  
 P88975 PRELIMINARY: PRT; 149 AA.  
 ID P88975;  
 AC P88975;  
 DT 01-MAY-1997 (TRENBLrel. 03, Created)  
 DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)

DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)  
 DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).  
 OS Macrotellium golden mosaic geminivirus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=51676;  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-JAMAICA;  
 RA Roye M.E.,  
 RI Thesis (1996), biochemistry, University of the West Indies, Jamaica.  
 DR EMBL: U67526; AAB57865.1; -  
 DR INTERPRO: IPR001191; -  
 DR INTERPRO: IPR001191; -  
 DR PFAM: PF00799; Gemin1\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATL1.  
 FT NON\_TER 149 149  
 SQ SEQUENCE 149 AA; 16785 MW; 84CF5ED4C9CD508 CRC64;

Query Match 73.5% Score 266; DB 12; Length 149;  
 Best Local Similarity 68.6% Pred. No. 1.7e-21;  
 Matches 48; Conservative 13; Mismatches 9; Indels 0; Gaps 0;  
 Qy 1 TUVGFEVQGAAGCGCOTNDAAEAALNASKEEALQIREKIPKYLQFPHNLSNU. 60  
 Db 52 TIVGFEVQDGSANGCGQTNDAAEAALNASKEEALQIREKIPKYLQFPHNLSNU. 111  
 Qy 61 DRFDKTP 70  
 Db 112 DRFMDPPEP 121

RESULT 8  
 Q9YL44 PRELIMINARY: PRT; 233 AA.  
 ID Q9YL44;  
 AC Q9YL44;  
 DT 01-MAY-1999 (TRENBLrel. 10, Created)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
 DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)  
 DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).  
 OS Macrotellium golden mosaic geminivirus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=51676;  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-JAMAICA STRAIN 1;  
 RA Roye M.E.,  
 RI Molecular characterization of two distinct geminiviruses infecting M.  
 alalytoides in Jamaica. (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF089940; AAD17850.1; -  
 DR INTERPRO: IPR001191; -  
 DR INTERPRO: IPR001191; -  
 DR PFAM: PF00799; Gemin1\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATL1.  
 SQ SEQUENCE 233 AA; 26355 MW; AA490AF4D2156A02 CRC64;  
 Query Match 73.5% Score 266; DB 12; Length 233;  
 Best Local Similarity 68.6% Pred. No. 2.9e-21;

Matches	48:	Conservative	13:	Mismatches	9:	Indels	0:	Caps	0:
QY	1	TLVNGEPQVDAAGAGCGQTSNDAAAEALNASSKEALQIIRKIPKYLPOFHNISSNL	60						
DB	110	TLVNGEPQVQIDGRARGGQQTSDNAAEALNSGTEKAAIRVKEIKPEKFLFOYHNLSNL	169						
QY	61	DRFDKTPPEP 70							
DB	170	DRLEKMOKEP 179							
RESULT	9								
Q39180									
ID	Q39180	PRELIMINARY:	PRY:	234	AA.				
DT	01-JAN-1998	(TREMblrel. 05, Created)							
DT	01-JAN-1998	(TREMblrel. 05, Last sequence update)							
DT	01-OCT-2000	(TREMblrel. 15, Last annotation update)							
QZ	REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).								
QZ	REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).								
QC	VIRUSES; AS2DNA VIRUSES; Geminiviridae; Begomovirus.								
QC	VIRUSES; AS2DNA VIRUSES; Geminiviridae; Begomovirus.								
NCBI_TaxId=10827;									
[1]									
RP	SEQUENCE FROM N.A.								
RC	STRATN-TOMATO STRAIN;								
RC	Guzman P., Arredondo C.R., Emmatty D., Portillo R.J., Gilbertson R.L.;								
RL	Plant Dis. 81:312-312(1997).								
RL	EMBL: AF026553; AB982605.1; .								
DR	INTERPRO: IPRO0131; .								
DR	IPR001310; .								
DR	PPAK: PF007759; GenInfo: Ali: 1.								
DR	PRINTS: PR00227; GEMCOATLVL.								
DR	PRINTS: PR00228; GEMCOATLVL.								
DR	NON_TER 234								
FT	SEQUENCE 234 AA; 26486 MW; 9ED8F0697105CD19 CRC64;								

	Query Match	Best Local Similarity	71.8%	Scores	260	DB 13	Length	234
	Matches	47	Conservative	14	Mismatches	9	Indels	0
QY	1	TLVNGFEYDVGAAAGGCQTSDNAAEALNASKEALQIREKIPKLYLQFHNINSL	60					
DB	110	TLNGLQFDIGRSGAGGQQTANDAAEALNSGTKEAKRIKPEKLPKLYQVHNLSDDL	169					
QY	61	DRTEKTPRP 70						
DB	170	DRINKAPNP 179						

```

RESULT 10
OSZ089
ID 09Z089 PRELIMINARY: ERT: 190 AA.
AC 09Z089
DT 01-MAY-1999 (TReMBRel. 10. Created)
DT 01-MAY-1999 (TReMBRel. 10. Last sequence update)
DT 01-MAY-1999 (TReMBRel. 10. Last annotation update)
DT C1 and C2 GENES: VOKOHAMA3;
DT C1 and C2 GENES: VOKOHAMA3;
DT PARTIAL AND COMPLETE CDS (FRAGMENT).
DT C1.
OS tobacco leaf curl virus.
OS OS viruses: ssDNA viruses: Geminiviridae: Begomovirus.
OC NCBI_TaxID=57762;
OC VITUS.
ID NCBI_TaxID=57762;
AC VITUS.
DT SEQUENCE FROM N.A.
DT STRAIN=VOKOHAMA3.
RC Ooi K., Ohshita S., Ishii I., Yahara T.:
RT "Molecular phylogeny of geminivirus infecting wild plants in Japan.":
RL J. Plant Res. 110:247-257 (1997).
RL EMBL:AA000130.1; BAA34033.1;
DR J. Plant Res. 110:247-257 (1997).
DR INTERPRO: IPR001301;
DR PFAM: PF00759; GeminA_L1.
DR

```

DR	PRINTS:	PH00227;	GEMCOATALLI.
DR	PRINTS:	PR00228;	GEMCOATLVLII.
FT	NON_TER	1	
FT	NON_TER	190	190
SQ	SEQUENCE	190 AA; 21433 MW; AAC093DID1610FAD CRC64;	

  

		Query Match	71.0%; Score 257; DB 12; Length 190;
		Best Local Similarity	58.8%; Pred. No. 2.2e-20;
		Matches	50; Conservative 9; Mismatches 10; Indels 16; Gaps 1;

  

QY	1	TLVNGEFOVDGAAAGGCQTSDNDAAREALNASKEEAQLITREIKPEKYLFOFHNSNI	60
DB	95	TLENGTFQIDRGAGGGCGNQANDCAELANASRAKLAIIIREKLPKDIFOYHNLSNI	144
QY	61	DRI--	69
QY	111		111
QY	145	DRIEAFPLEVFCPTTASSFDQVPE	169

  

RESULT	11	
QZ0208	C	PRELIMINARY; PRT: 190 AA.
AC	092094	
AD	092084	
DT	01-MAY-1999	(TREMBREL. 10. Created)
DT	01-MAY-1999	(TREMBREL. 10. Last sequence update)
DT	01-MAY-2000	(TREMBREL. 13. Last annotation update)
DE	C1 AND C4 GENES, CLONE YOKOHAMAS-2,	
DE	PARTIAL AND COMPLETE CDS (FRAGMENT).	
ON	Tabacco leaf curl virus.	
QC	Viruses; ssDNA viruses, Gemniviiridae; Begomovirus.	
OC	NCBI_TaxID=67762;	
RN	[1]	
RX	SEQUENCE FROM N.A.	
RP	STRAIN-YOKOHAMAS;	
RC	Ooi K., Oshita S., Ishii I., Yahara T.;	
RA	A molecular phylogeny of geminiviruses infecting wild plants in Japan.";	
RT	Molecular Biology and Evolution 1999;16(12):1997.	
RM	EMBL: AB001138; BAA34039.1;	
DR	INTERPRO: IPRO01191;	
DR	INTERPRO: IPRO01301;	
DR	PFAM: PF00799; Gemini_L1;	
DR	PRINTS: PR00227; GEMCOATALLI.	
DR	PRINTS: PR00228; GEMCOATLVLII.	
FT	NON_TER	1
FT	NON_TER	190
SQ	SEQUENCE	190 AA; 21444 MW; AACIC2943EF3EQIAD CRC64;

	Query Match Pair Local	Similarity 58.9%	Score 267	DB 12	Length 190;
	Matches 50;	Conservative 9;	Mismatches 10;	Indels 16;	Gaps 1;
Qy	1 TLWGSEFQVDAAGACGCT	SNDAAAEALANASKEEALQITREKIPKTLYTFPHNLNSLN	60		
Dy	45 TLEWTEFTQIDGRSAGCGGNANDCAELANAKRAELAIITREKPKDFIYHNLSN	L 144			
Qy	61 DRT	---FDKTPGE	69		
Dy	145 DRFEAPLEVFCPTASSFDDQVE	169			
RESULT 12					
ID Q94827					
AC Q94827	PRELIMINARY;	EPR:	190 AA.		
DT 01-NOV-1999	(TREMBLrel. 12. Created)				
DT 01-MAY-1999	(TREMBLrel. 12. Last sequence update)				
DT 01-MAY-2000	(TREMBLrel. 13. Last annotation update)				
GN C1	GF PROTEIN (FRAGMENT).				
OS Tobacco leaf curl virus.					

OC Viruses: ssDNA viruses: Geminiviridae; Begomovirus.

```

RN NCBI_TaxID=67762;
RC STRAIN=ABUR3;
RA Ool K., Ohshita S., Ishii I., Yahara T.;
RT *Molecular phylogeny of geminivirus infecting wild plants in Japan.*;
RL J. Plant Res. 110:247-257(1997).
DR EMBL: AB001303; BAA34010.1; -.
DR INTERPRO: IPR001191; -.
DR PFAM: PF00799; Gemin1_1; 1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00228; GEMCOATL1.
FT NON_TER 1 190
FT SEQUENCE 190 AA: 21444 MW: 93C3742ABRDB7ER CRC64;
SQ

```

```

Query Match 70.78: Score 256; DB 12; Length 190;
Best Local Similarity 58.84; Pred. No. 2.8e-20;
Matches 50; Conservative 9; Mismatches 10; Indels 16; Gaps 1;

Oy 1 TLVWGSEFVQCGAARAGCCGTNDAAAEALNASSKEEALQIIEKPKYLFQVHNLSNL 60
Db 85 TLEWGTFOVGRSANGCCONADACAEALNASSKAELSIIIEKLPKDFIFQVHNLSNL 144
Oy 61 DRI-----FQKTPK 69
Db 145 DRIAPFLEVPFCPSNSSFQVPE 169

```

```

RESULT 13
Oy2004 PRELIMINARY: PRT: 208 AA.
ID Q920C4
AC Q920B8
DT 01-MAY-1999 (TEMBLrel. 10, Created)
DT 01-MAY-1999 (TEMBLrel. 10, Last sequence update)
DT 01-MAY-2000 (TEMBLrel. 13, Last annotation update)
DE C1 AND C4 GENES; CLONE ABUR3-1, PARTIAL
DE PARTIAL AND COMPLETE CDS (FRAGMENT).
GN C1.
GN C4.
GN Tobacco leaf curl virus.
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=67762;
RN STRAIN=ABUR3;
RP SEQUENCE FROM N.A.

```

```

RC Ool K., Ohshita S., Ishii I., Yahara T.;
RT *Molecular phylogeny of geminivirus infecting wild plants in Japan.*;
RL J. Plant Res. 110:247-257(1997).
DR EMBL: AB001294; BAA33992.1; -.
DR INTERPRO: IPR001191; -.
DR PFAM: PF00799; Gemin1_1; 1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00228; GEMCOATL1.
FT NON_TER 208 208
FT SEQUENCE 208 AA: 23526 MW: 249CC1DB729C72D CRC64;
SQ

```

```

Query Match 70.78: Score 256; DB 12; Length 208;
Best Local Similarity 58.84; Pred. No. 3.1e-20;
Matches 50; Conservative 9; Mismatches 10; Indels 16; Gaps 1;

Oy 1 TLVWGSEFVQCGAARAGCCGTNDAAAEALNASSKEEALQIIEKPKYLFQVHNLSNL 60
Db 102 TLEWGTFOVGRSANGCCONADACAEALNASSKAELSIIIEKLPKDFIFQVHNLSNL 161
Oy 61 DRI-----FQKTPK 69
Db 162 DRIAPFLEVPFCPSNSSFQVPE 186

```

```

RESULT 14
Oy2004 PRELIMINARY: PRT: 208 AA.
ID Q920C4
AC Q920B8
DT 01-MAY-1999 (TEMBLrel. 10, Created)
DT 01-MAY-1999 (TEMBLrel. 10, Last sequence update)
DT 01-MAY-2000 (TEMBLrel. 13, Last annotation update)
DE C1 AND C4 GENES; CLONE AMG-1(B152),
DE PARTIAL AND COMPLETE CDS (FRAGMENT).
GN C1.
GN C4.
GN Tobacco leaf curl virus.
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=67762;
RN STRAIN=AMGB152;
RP SEQUENCE FROM N.A.
RC Ool K., Ohshita S., Ishii I., Yahara T.;
RT *Molecular phylogeny of geminivirus infecting wild plants in Japan.*;
RL J. Plant Res. 110:247-257(1997).
DR EMBL: AB001296; BAA33996.1; -.
DR INTERPRO: IPR001191; -.
DR PFAM: PF00799; Gemin1_1; 1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00228; GEMCOATL1.
FT NON_TER 208 208
FT SEQUENCE 208 AA: 23486 MW: B301135F799C30AD CRC64;
SQ

```

```

Query Match 70.28: Score 254; DB 12; Length 208;
Best Local Similarity 57.64; Pred. No. 5.1e-20;
Matches 49; Conservative 10; Mismatches 10; Indels 16; Gaps 1;

Oy 1 TLVWGSEFVQCGAARAGCCGTNDAAAEALNASSKEEALQIIEKPKYLFQVHNLSNL 60
Db 102 TLEWGTFOVGRSANGCCONADACAEALNASSKAELSIIIEKLPKDFIFQVHNLSNL 161
Oy 61 DRI-----FQKTPK 69
Db 162 DRIAPFLEVPFCPSNSSFQVPE 186

```

```

RESULT 15
Oy20B8 PRELIMINARY: PRT: 208 AA.
ID Q920B8
AC Q920B8
DT 01-MAY-1999 (TEMBLrel. 10, Created)
DT 01-MAY-1999 (TEMBLrel. 10, Last sequence update)
DT 01-MAY-2000 (TEMBLrel. 13, Last annotation update)
DE C1 AND C4 GENES; CLONE AMG-2(B154),
DE PARTIAL AND COMPLETE CDS (FRAGMENT).
GN C1.
GN C4.
GN Tobacco leaf curl virus.
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=67762;
RN STRAIN=AMGB154;
RP SEQUENCE FROM N.A.
RC Ool K., Ohshita S., Ishii I., Yahara T.;
RT *Molecular phylogeny of geminivirus infecting wild plants in Japan.*;
RL J. Plant Res. 110:247-257(1997).
DR EMBL: AB001297; BAA33998.1; -.
DR INTERPRO: IPR001191; -.
DR PFAM: PF00799; Gemin1_1; 1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00228; GEMCOATL1.
FT NON_TER 208 208
FT SEQUENCE 208 AA: 23472 MW: 62900DE7C9956AA CRC64;
SQ

```

Query Match 70.2%; Score 254; DB 12; Length 208;  
Best Local Similarity 57.6%; Percent Identical 54.8%;  
Matches 49; Conservative 10; Mismatches 10; Indels 16; Gaps 1;  
QY 1 TLVWGEFQVDGAAAGGCGTSDNDAAEALNASSKEEALQIIREKIPEKYLQFHNLSNL 60  
DB 102 TLEWGTQIDGHSARGGCONADAEALNASSKADALAIIREKLEKDFIYHNLNSNL 161  
QY 61 DRI-----FDKYPE 69  
DB 162 DRIFAPPLEVFVCPFSSSFDQYPE 186

Search completed: February 3, 2001, 02:22:50  
Job time: 1887 sec

GenCode version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.  
OM protein - protein search, using sw model  
Run on: February 3, 2001, 02:24:24 ; Search time 83.07 Seconds  
(without alignments)  
27.213 Million cell updates/sec

Title: US-09-289-346a-2  
Perfect score: 362  
Sequence: 1 TLVWGEFQVQGAAGSGCOT.....FOFHNLNSLDRFDKTEPP 70

Scoring Table: BLOSUM62  
Gapop 10 0 ; Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DP seq length: 0  
Maximum DP seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	349	96.4	352	1 VAL1_TGMV	P03567 tomato gold
2	249	68.8	361	1 VAL1_PYMV	P27258 potato yell
3	236	62.2	358	1 VAL1_CIVR	P14992 cassava lat
4	236	62.2	358	1 VAL1_CIVR	P14992 cassava lat
5	229	61.3	352	1 VAL1_TVLC	P36276 cassava yell
6	229	61.3	352	1 VAL1_TVLC	P36276 cassava yell
7	218	60.2	349	1 VAL1_PTHV	P38609 tomato yell
8	215	59.4	359	1 VAL1_TVLC	Q06923 pepper huas
9	213	58.6	353	1 VAL1_BGVW	P27260 tomato yell
10	212	58.6	353	1 VAL1_BGVW	P05175 bean golden
11	209	57.7	358	1 VAL1_RCTV	P14991 beet curly
12	209	57.7	358	1 VAL1_RCTV	P14991 beet curly
13	200	55.2	357	1 VAL1_TVLC	P27259 tomato yell
14	124	34.3	347	1 VAL1_SICV	P29048 squash leaf
15	68.5	18.9	1713	1 LMA3_HUMAN	Q16787 homo sapien
16	65.5	18.1	1610	1 CCAD_MESAU	Q09244 mesocricetu
17	65.5	18.1	1610	1 CCAD_HUMAN	Q01668 homo sapien
18	65.5	18.1	2203	1 CCAD_RAT	P27332 rattus norv
19	64.5	17.8	259	1 Y175_HELHPJ	Q26072 hellicobacte
20	64.5	17.8	259	1 Y175_HELHPJ	Q26072 hellicobacte
21	64.5	17.8	335	1 G3P_BACCO	P15115 bacillus at
22	64.5	17.8	557	1 SASB_ANAPL	Q04791 anas platyr
23	62.5	17.3	419	1 Y223_MYCN	P75465 mycobasma
24	62	17.1	285	1 VAL1_TVDVA	P31617 tobacco yell
25	62	17.1	1483	1 CYP1_YEAST	P12351 saccharomyc
26	60.5	16.7	136	1 Y452_CAREPL	O62250 caenorhabdi
27	60	16.6	407	1 CUS1_DPOME	P26072 human papill
28	60	16.6	407	1 CUS1_DPOME	P26072 human papill
29	60	16.6	513	1 HKM0_CHICK	P18080 gallus gall
30	59	16.3	129	1 RK12_PORPU	P51339 porphyra pu
31	59	16.3	247	1 YCF4_YEAST	P25349 saccharomyc
32	58.5	16.2	299	1 Y175_HELHPJ	P56112 hellicobacte
33	58.5	16.2	589	1 YB67_HAEIN	P45221 haemophilus

34 58.5 16.2 2190 1 CCAD\_CHICK  
35 58 16.0 387 1 Y4PF\_RHISN  
36 58 16.0 617 1 YACH\_ECOLI  
37 58 16.0 874 1 SLAP\_BACLI  
38 57.5 15.9 266 1 Y223\_MYCN  
39 57.5 15.9 266 1 Y223\_MYCN  
40 57.5 15.9 525 1 P2BB\_RAT  
41 57.5 15.9 1852 1 CCAS\_CYPKA  
42 57 15.7 312 1 SOX3\_CHICK  
43 57 15.7 351 1 CRTB\_RHOS  
44 57 15.7 382 1 IPAC\_SHIDY  
45 57 15.7 393 1 METK\_ASCIM

073700 gallus gall  
P55615 rhizobium s  
P36882 escherichia  
P49052 bacillus li  
P33075 mycobacteri  
O33927  
P20551 rattus norv  
P22316 cyprinus ca  
P48433 gallus gall  
P54905 rhodobacter  
Q03946 shigella dy  
P50304 ascobolus i

ALIGNMENTS

RESULT 1  
VAL1\_TGMV STANDARD: PRT: 352 AA.  
ID VAL1\_TGMV  
AC  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE ALL PROTEIN.  
GN ACL.  
OS Tomato golden mosaic virus (TGMV).  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
RN  
RP SEQUENCE FROM N.A.  
EA Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.:  
RT "Complete nucleotide sequence of the infectious cloned DNA components  
of tomato golden mosaic virus: potential coding regions and regulatory  
sequences." ;  
RL EMBD J. 3:2197-2205(1984).  
CC -- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/commence/  
or send an email to license@isb-sib.ch).  
CC EMBL: K02029; ; NOT\_ANNOTATED\_CDS.  
DR PIR: A04170; OOCVLI  
DR INTERPRO: IPR001191; ;  
DR INTERPRO: IPR001301; ;  
DR PFAM: PF00759; Gemin1-AL1; 1.  
DR PRINTS: PR00227; GEMCONTALI.  
DR PROSITE: PR00228; GEMCONTCLVL.  
KW ATP-binding, 223 230  
FT NP\_BIND 352 AA; 40332 MW; C33C938S9564B4A4 CRC64;  
SQ SEQUENCE 352 AA; 40332 MW; C33C938S9564B4A4 CRC64;

Query Match 96.4%; Score 349; DB 1; Length 352;  
Best Local Similarity 95.7%; Pred. No. 8-31;  
Matches 67; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 TLVWGEFQVQGAAGSGCOTSNDAARALNASKKEALQITREKPKYLFQFHNLNSL 60  
Db 111 TLVWGEFQVQGRSGGCGTSDAAARALNASKKEALQITREKPKYLFQFHNLNSL 170  
Oy 61 DRFDKTEPP 70  
Db 171 DRFDKTEPP 180

RESULT 2  
VAL1\_PYMV STANDARD: PRT: 361 AA.

AC DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DE ALL PROTEIN (Rel. 23, Last annotation update)  
 DE ALL PROTEIN (Rel. 23, Last annotation update)  
 OS Potato yellow mosaic virus (isolate Venezuela).  
 OS Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-91311403; PubMed-1856690;  
 RA Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.:  
 RT The nucleotide sequence of the infectious cloned DNA components of  
 RT Potato yellow mosaic virus.  
 RL J. Gen. Virol. 72:1515-1520(1991).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: D00940; BAA00782.1; -  
 CC PIR: J00364; Q0CVPT.  
 CC INTERPRO: IPR001191; -  
 CC PRFAM: PF00759; Gemin1.AL1; 1.  
 CC PRINTS: PR00227; GEMCOATL1.  
 CC PRINITS: PR00228; GEMCOATCLV1.  
 KW ATP-binding.  
 FT NP\_BIND 222 229  
 FT SEQUENCE 361 AA; 40950 MW; 5627A33BF1264383 CRC64;  
 SQ

Query Match 58.8%; Score 249; DB 1; Length 361;  
 Best Local Similarity 56.7%; Pred. No. 6; 6e-20;  
 Matches 46; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDAAGAGCGTSDNDAALNASSKEALQIIREKIPKYLQFHNLSNL 60  
 DB 110 TLVWGEFQVDAAGAGCGTSDNDAALNASSKEALQIIREKIPKYLQFHNLSNL 169  
 QY 61 DRIFDKPTE 69  
 DB 170 DRIFDKAPE 178

RESULT 3  
 ID VALL CLVW STANDARD; PRT: 358 AA.  
 AC P14382;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DE ALL PROTEIN (Rel. 29, Last annotation update)  
 DE ALL PROTEIN (40.4 KDA PROTEIN).  
 OS Cassava latent virus (strain West Koyan 844).  
 OS Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Stanley J., Gay M.R.:  
 RT Nucleotide sequence of cassava latent virus DNA. \*;  
 RL Nucleic Acids Res. 10:260-262(1983).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: J02057; -; NOT\_ANNOTATED\_CDS.  
 CC INTERPRO: IPR001191; -  
 CC PRFAM: PF00759; Gemin1.AL1; 1.  
 CC PRINTS: PR00227; GEMCOATL1.  
 CC PRINITS: PR00228; GEMCOATCLV1.  
 KW ATP-binding.  
 FT NP\_BIND 220 227  
 FT SEQUENCE 358 AA; 40346 MW; ED173E753EE92D69 CRC64;  
 SQ

Query Match 65.2%; Score 236; DB 1; Length 358;  
 Best Local Similarity 60.0%; Pred. No. 1; 7e-18;  
 Matches 42; Conservative 14; Mismatches 14; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDAAGAGCGTSDNDAALNASSKEALQIIREKIPKYLQFHNLSNL 60  
 DB 109 TLVWGEFQVDAAGAGCGTSDNDAALNASSKEALQIIREKIPKYLQFHNLSNL 168  
 QY 61 DRIFDKPTE 70  
 DB 169 DRIFQEPAP 178

RESULT 4  
 ID VALL CLVW STANDARD; PRT: 358 AA.  
 AC P14372;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DE ALL PROTEIN (40.4 KDA PROTEIN).  
 DE ALL PROTEIN (40.4 KDA PROTEIN).  
 GN ACI.  
 OS Cassava latent virus (strain Nigerian).  
 OS Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-90174930; PubMed-2308831;  
 RA Morris B., Coates L., Lowe S., Richardson K., Eddy P.:  
 RT Nucleotide sequence of the infectious cloned DNA components of  
 RT African cassava mosaic virus (Nigerian strain). \*;  
 RL Nucleic Acids Res. 17:103-107(1989).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X17095; CAA34953.1; -  
 CC PIR: S07594; S07594.  
 CC INTERPRO: IPR001191; -  
 CC PRFAM: PF001301; -  
 CC PRINTS: PR00227; GEMCOATL1.  
 CC PRINITS: PR00228; GEMCOATCLV1.  
 KW ATP-binding.  
 FT NP\_BIND 220 227  
 FT SEQUENCE 358 AA; 40435 MW; 1DB16B80CB2D52C CRC64;  
 SQ

Query Match 65.2%; Score 236; DB 1; Length 358;  
 Best Local Similarity 60.0%; Pred. No. 1; 7e-18;  
 Matches 42; Conservative 14; Mismatches 14; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDAAGAGCGTSDNDAALNASSKEALQIIREKIPKYLQFHNLSNL 60  
 DB 109 TLVWGEFQVDAAGAGCGTSDNDAALNASSKEALQIIREKIPKYLQFHNLSNL 168  
 QY 61 DRIFDKPTE 70  
 DB 169 DRIFQEPAP 178

RESULT 4  
 ID VALL CLVW STANDARD; PRT: 358 AA.  
 AC P14372;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DE ALL PROTEIN (40.4 KDA PROTEIN).  
 DE ALL PROTEIN (40.4 KDA PROTEIN).  
 GN ACI.  
 OS Cassava latent virus (strain Nigerian).  
 OS Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-90174930; PubMed-2308831;  
 RA Morris B., Coates L., Lowe S., Richardson K., Eddy P.:  
 RT Nucleotide sequence of the infectious cloned DNA components of  
 RT African cassava mosaic virus (Nigerian strain). \*;  
 RL Nucleic Acids Res. 17:103-107(1989).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----



```

QY 61 DRIFDKTPEP 70
Db 169 DRIFQPPAP 178

RESULT 5
VALL_TYLCU STANDARD; PRT; 362 AA.
AC P36279;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE ALL PROTEIN (C1 PROTEIN).
OS Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN
RP SEQUENCE FROM N.A. PubMed=8423446;
RA DRYB, Riden J, Kraker R, Mullineaux P M, Rezaian M A.;
RT "Nucleotide sequence and genome organization of tomato leaf curl
geminivirus.";
RL J. Gen. Virol. 74:147-151(1993).
CC -1. SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL; Z53251; CAB50886.1; -.
DR PIR: JQ2300;
DR INTERPRO: IPR001191; -.
DR PFAM: PF00799; Gemin_L1;
DR PRINTS: PR00227; GEMCOATL1.
KW ATP-binding, 221 228 ATP (POTENTIAL).
FT NP_BIND
SQ SEQUENCE 362 AA; 41197 MW; 343E7184B4704098 CRC64;

Query Match 63.3%; Score 229; DB 1; Length 362;
Best Local Similarity 52.9%; Pred. No. 9.8e-18;
Matches 45; Conservative 10; Mismatches 14; Indels 16; Gaps 1;

QY 1 TLVGEFQVDAAGAGCGTSDNDAAEALNASSKEALQIIRKIPKYLQFHNLSNL 60
Db 110 TLVGEFQIDGRSARGGQSDANVAQLNKGSKSEALNVLRELAPDYVLOFHNLSNL 169

QY 61 DRI-----FKTPE 69
Db 170 DRIFTPLEVTVSPFLSSFDKRVPE 194

RESULT 6
VALL_TYLCU STANDARD; PRT; 359 AA.
AC P38609;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE ALL PROTEIN (C1 PROTEIN).
OS Tomato yellow leaf curl virus (strain Murcia) (TYLCV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN
RP SEQUENCE FROM N.A.
RA MEDLINE=94256836; PubMed=8198442;
RX INTERPRO: IPR001191; -.

Noris E., Hidalgo E., Accotto G., Moriones E.;
"High similarity among the tomato yellow leaf curl virus isolates
from the west Mediterranean basin: the nucleotide sequence of an
infectious clone from Spain";
Arch. Virol. 135:163-170(1994).
CC SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL; Z25751; CAB81026.1; -.
DR PIR: S39211; S39211.
DR INTERPRO: IPR001191; -.
DR PFAM: PF00799; Gemin_L1;
DR PRINTS: PR00227; GEMCOATL1.
KW ATP-binding, 221 228 ATP (POTENTIAL).
FT NP_BIND
SQ SEQUENCE 359 AA; 41065 MW; 2D170A51EF80A3BC CRC64;

Query Match 51.3%; Score 222; DB 1; Length 359;
Best Local Similarity 38.0%; Pred. No. 5.6e-17;
Matches 40; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

QY 2 LVKGEFQVDAAGAGCGTSDNDAAEALNASSKEALQIIRKIPKYLQFHNLSNL 61
Db 111 LVKGEFQIDGRSARGGQSDANDATKATNAGSKSEALDKELAPDYVLOFHNLSNL 170

QY 62 DRIFDKTPEP 70
Db 171 DRIFQPPAP 179

RESULT 7
VALL_PHVU STANDARD; PRT; 349 AA.
ID VALL_PHVU
AC P34909;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE ALL PROTEIN.
GN ALL.
OS Pepper huasteco virus (PHV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN
RP SEQUENCE FROM N.A.
RA MEDLINE=94015007; PubMed=8409944;
RA Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,
Riveka-Bustamante R.F.;
RT "Complete nucleotide sequence of pepper huasteco virus: analysis and
comparison with bipartite geminiviruses.";
RL J. Gen. Virol. 74:2225-2231(1993).
CC -1. SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL; X70418; CAB49856.1; -.
DR PIR: S31875; S31875.
DR PIR: JQ2300; JQ2300.
DR INTERPRO: IPR001191; -.

```

[illegible]

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is not removed, its  
 CC modified and this statement is not removed, usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X15983; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL: X15983; -; NOT\_ANNOTATED\_CDS.  
 DR INTERPRO: IP0001191; -;  
 DR IP0001191; -;  
 DR PFAM: PF00759; Gemini\_A1; 1.  
 DR PRINTS: PR00227; GEMC0ATL1.  
 DR PRINTS: PR00228; GEMC0ATL1.  
 DR ATP-binding. 221 228 ATP (POTENTIAL).  
 KW NP-BIND 221 228  
 FT SEQUENCE 355 AA; 40257 MW; 16A2CABA63251E95 CRC64;

Query Match 58.6%; Score 212; DB 1; Length 355;  
 Best Local Similarity 55.7%; Pred. No. 6.8e-16;  
 Matches 39; Conservative 13; Mismatches 18; Indels 0; Gaps 0;

Qy 1 TLWGEFVQDGAAGCGCTNDAAAEALNASSKEALQIIRKIPKYLQFHLNLSN 60  
 Db 110 TAENGFEQIDGARGCGQSTANAKALNMGVQSALNKEQPKDYVLONHINSNL 169  
 Qy 61 DRIFDKTPEP 70  
 Db 170 ERIKFAPEP 179

RESULT 11  
 VAL1\_BCTV  
 ID VAL1\_BCTV STANDARD; PRT; 358 AA.  
 AC P14991;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 01-AUG-1992 (Rel. 23, Last annotation update)  
 DE AL1 PROTEIN (40.8 KDA PROTEIN).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is not removed, its  
 CC modified and this statement is not removed, usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X04144; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL: X04144; -; NOT\_ANNOTATED\_CDS.  
 DR INTERPRO: IP0001191; -;  
 DR IP0001191; -;  
 DR PFAM: PF00759; Gemini\_A1; 1.  
 DR PRINTS: PR00227; GEMC0ATL1.  
 DR PRINTS: PR00228; GEMC0ATL1.  
 DR ATP-binding. 222 229 ATP (POTENTIAL).  
 KW NP-BIND 222 229  
 FT SEQUENCE 358 AA; 40889 MW; 39A45FE3C0B9C333 CRC64;

Query Match 57.7%; Score 209; DB 1; Length 358;  
 Best Local Similarity 52.9%; Pred. No. 1.4e-15;  
 Matches 37; Conservative 17; Mismatches 16; Indels 0; Gaps 0;

Matches 37; Conservative 17; Mismatches 16; Indels 0; Gaps 0;  
 Qy 1 TLWGEFVQDGAAGCGCTNDAAAEALNASSKEALQIIRKIPKYLQFHLNLSN 60  
 Db 110 TLWGEFVQDGAAGCGCTNDAAAEALNASSKEALQIIRKIPKYLQFHLNLSN 169  
 Qy 61 DRIFDKTPEP 70  
 Db 170 QRIKFAPEP 179

RESULT 12  
 VAL1\_TMOV  
 ID VAL1\_TMOV STANDARD; PRT; 361 AA.  
 AC Q06657;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE AL1 PROTEIN.

OS Tomato mottle virus (isolate Florida) (TMov).  
 OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-93107858; PubMed-1469361;  
 RA Abouzid A.M., Polston J.E., Hiebert E.;  
 RT Isolation and characterization of tomato mottle virus, a new geminivirus  
 BL J. Gen. Virol. 73:3225-3229 (1992).

CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is not removed, its  
 CC modified and this statement is not removed, usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: L14450; AAC32414.1; -;  
 DR PIR: JQ1870; JQ1870.  
 DR INTERPRO: IP0001191; -;  
 DR IP0001191; -;  
 DR PFAM: PF00759; Gemini\_A1; 1.  
 DR PRINTS: PR00227; GEMC0ATL1.  
 DR PRINTS: PR00228; GEMC0ATL1.  
 KW NP-binding.  
 FT NP-BIND 222 229 ATP (BY SIMILARITY).  
 S0 SEQUENCE 361 AA; 40516 MW; 8138B65CEAC6950 CRC64;

Query Match 57.7%; Score 209; DB 1; Length 361;  
 Best Local Similarity 52.9%; Pred. No. 1.5e-15;  
 Matches 37; Conservative 17; Mismatches 16; Indels 0; Gaps 0;

Qy 1 TLWGEFVQDGAAGCGCTNDAAAEALNASSKEALQIIRKIPKYLQFHLNLSN 60  
 Db 110 TLWGEFVQDGAAGCGCTNDAAAEALNASSKEALQIIRKIPKYLQFHLNLSN 169  
 Qy 61 DRIFDKTPEP 70  
 Db 170 ERIKFAPEP 179

RESULT 13  
 VAL1\_TYLCV  
 ID VAL1\_TYLCV STANDARD; PRT; 357 AA.  
 AC P27259;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE AL1 PROTEIN (C1 PROTEIN).

OS Tomato yellow leaf curl virus (TYLCV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 RN  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE-92024070; PubMed-1926771;  
 RA Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.:  
 RT "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus  
 RL with a single genomic component."; <http://www.isb-sib.ch/announce/>  
 RL Virology 185:151-161(1991).  
 CC  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 CC at European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: X15556; CAB33688.1;  
 CC PIR: D40779; Q0CVG3;  
 CC DR INTERPRO: IPR001301;  
 CC DR PFAM: PF00799; Gemini\_AL1;  
 CC DR PRINTS: PR00227; GEMCOATALL.  
 CC DR PRINTS: PR00228; GEMCOATCLV1.  
 CC KW ATP-binding; 319 226 ATP (POTENTIAL).  
 CC FT NP\_BIND 319 226  
 CC SEQUENCE 357 AA; 40578 MW; 933AB68E1A3B2A7 CRC64;  
 CC  
 CC Query Match 55.2%; Score 200; DB 1; Length 357;  
 CC Best local similarity 62.3%; Pred. No. 1.4e-14;  
 CC Matches 38; Conservative 11; Mismatches 12; Indels 0; Gaps 0;  
 CC  
 CC QY 4 WGFQVGVGAAGAGCCGTSNDAAEAALNASKEALQIREKIPKLYFOFNLSNLDRI 63  
 CC DB 111 FVSOLDGRSARGSQSNDAAEAALNASGSKSEALNILKERAPKPYILOFNLSNLDRI 170  
 CC  
 CC QY 64 F 64  
 CC DB 171 F 171  
 CC  
 CC RESULT 14  
 CC VAL1\_SICV STANDARD: PRT: 347 AA.  
 CC AC P29048;  
 CC DT 01-DEC-1992 (Rel. 24, Created)  
 CC DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 CC DE ALI PROTEIN (Rel. 24, Last annotation update)  
 CC OS Squash leaf curl virus.  
 CC RN Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RA Lazzerotti S.G., Lazdinis I.B.: peptide sequences of the cloned genomic  
 CC components of a bipartite squash leaf curl geminivirus with a broad  
 CC host range phenotype."; <http://www.isb-sib.ch/announce/>  
 CC RL Virology 180:58-69(1991).  
 CC  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 CC at European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: M38163; AAC32410.1; ALT\_INIT.

DR PIR: C36785; Q0CVS1.  
 DR INTERPRO: IPR001191;  
 DR PRINTS: PR001301;  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR PRINTS: PR00228; GEMCOATCLV1.  
 KW ATP-binding.  
 FT NP\_BIND 218 225 ATP (POTENTIAL).  
 SQ SEQUENCE 347 AA; 39110 MW; AFDABDE122110E CRC64;  
 CC  
 CC Query Match 34.3%; Score 124; DB 1; Length 347;  
 CC Best local similarity 37.3%; Pred. No. 1.4e-06;  
 CC Matches 25; Conservative 14; Mismatches 23; Indels 1;  
 CC  
 CC QY 5 GEFQVGVGAAGAGCCGTSNDAAEAALNASKEALQIREKIPKLYFOFNLSNLDRI 64  
 CC DB 111 QVYKSSG----GSKSNKQDVYINAVNAGSAGEALDIKAGDKPTFVYINLLANVERLF 171  
 CC  
 CC QY 65 DKTPPE 70  
 CC DB 172 QKTPPE 177  
 CC  
 CC RESULT 15  
 CC LMA3\_HUMAN STANDARD: PRT: 1713 AA.  
 CC ID LMA3\_HUMAN  
 CC IC 019-NOV-1997 (Rel. 35, Created)  
 CC DT 01-NOV-1997 (Rel. 35; Last sequence update)  
 CC DT 01-OCT-2000 (Rel. 40; Last annotation update)  
 CC DE LAMININ ALPHA-3 CHAIN PRECURSOR (EPITILGRIN 170 KDA SUBUNIT) (E170).  
 CC GN LAMA3.  
 CC OS Homo sapiens (Human).  
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC RN [1] Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CC RP SEQUENCE FROM N.A.  
 CC RC TISSUE=KERATINOCYTES;  
 CC RX MEDLINE-94357926; PubMed-8077230;  
 CC RA Rynn M.C., Tizard R., Vandevanter D.R., Carter W.G.:  
 CC "Cloning of the LAMA3 gene encoding the alpha 3 chain of the adhesive  
 CC ligand epiligrin. Expression in wound repair."; <http://www.isb-sib.ch/announce/>  
 CC RL J Biol. Chem. 269:22779-22787(1994).  
 CC  
 CC RP SEQUENCE OF 1-1331 FROM N.A. (ISOFORMS A AND B).  
 CC RX MEDLINE-96163880; PubMed-8566427;  
 CC RA Vidal F., Baudouin C., Miquel C., Galliano M.-F., Chrisliano A.M.,  
 CC Utito J., Ortonne J.-P., Meneguizzi G.:  
 CC "Cloning of the laminin alpha 3 chain gene (LAMA3) and identification  
 CC of a homozygous deletion in a patient with Herlitz junctional  
 CC epidermolysis bullosa." <http://www.isb-sib.ch/announce/>  
 CC RL J Invest. Dermatol. 104:103-107(1995).  
 CC  
 CC -1- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ  
 CC IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF  
 CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING  
 CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.  
 CC -1- FUNCTION: LAMININ-5 IS THOUGHT TO BE INVOLVED IN (1) CELL ADHESION  
 CC AND (2) CELL MIGRATION. (2) SIGNAL TRANSDUCTION VIA FIBRONECTIN  
 CC RECEPTOR. (3) PHOSPHORYLATION OF PP125-PAK AND p80, (3) DIFFERENTIATION OF  
 CC KERATINOCYTES.  
 CC -1- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE  
 CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND  
 CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE  
 CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.  
 CC THE LONG ARM IS BETA, THE SHORT ARMS ARE ALPHA AND GAMMA. (K-LAMININ/  
 CC NICLIN) AND POSSIBLY ALSO A COMPONENT OF LAMININ-9 (K-LAMININ)  
 CC AND LAMININ-7 (K5-LAMININ).  
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT  
 CC MEMBRANES (MAJOR COMPONENT).  
 CC -1- ALTERNATIVE PRODUCTS: THE TWO ISOFORMS A AND B DIFFER IN THEIR N-  
 CC TERMINUS. THE SEQUENCE SHOWN HERE IS THAT OF THE SMALLER VARIANT

Search completed: February 3, 2001, 02:24:24  
Job time: 636 sec

---

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:17:36 ; Search time 118.74 Seconds  
40.029 Million cell updates/sec

Title: US-09-289-346a-2  
Perfect score: 362  
Sequence: 1 TLVGFQVQDGAAGAGCQT.....FQFHNLNLDLFDKTEP 70

Scoring table:  
BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0  
Maximum db seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 66:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description
1	349	96.4	352	1	QCVCVL	1	All protein - toma
2	249	68.8	361	1	QCVCPT	1	All protein - toma
3	236	65.2	358	2	S07594	2	hypothetical prote
4	229	63.3	362	1	J01887	1	All protein - toma
5	222	61.3	359	2	S39211	1	gene Cl protein -
6	218	60.2	349	2	J02490	1	replicase - pepper
7	215	59.4	359	2	S22593	1	All protein - pepp
8	215	59.4	359	2	S22593	1	All protein - pepp
9	212	58.6	351	2	J02327	1	All protein - indi
10	212	58.6	355	1	QCVCW1	1	AV1 protein - abut
11	209	57.7	358	1	J01870	1	All protein - toma
12	209	57.7	359	2	S39235	1	gene Cl protein -
13	209	57.7	385	2	S28360	1	All protein - beet
14	202	55.8	360	2	S59885	1	replication-associ
15	196	54.9	361	1	S45059	1	All protein - toma
16	142	39.2	331	2	S45059	1	All protein - toma
17	124	34.3	347	1	QCVCN1	1	All protein - squa
18	68.5	18.9	1713	2	A55347	1	adhesive ligand ep
19	65.5	18.1	1610	2	A46227	1	voltage-dependent
20	65.5	18.1	1646	2	JH0422	1	voltage-dependent
21	65.5	18.1	2161	2	JH0564	1	calcium channel al
22	65.5	18.1	2181	2	A38198	1	calcium channel al
23	65.5	18.1	2181	2	A38198	1	voltage-dependent
24	65.5	18.0	2371	2	B43502	1	probable peptidyl-
25	64.5	17.8	299	2	B71967	1	glycerol dehydratase
26	64.5	17.8	335	2	J50164	1	glycerol dehydratase
27	64.5	17.8	557	2	A47162	1	thiolesterase B (E
28	63.5	17.5	334	1	DBRSQF	1	glycerol dehydratase
29	63.5	17.5	392	2	T45290	1	acetyl-CoA C-acety

30 63 17.4 397 2 B71078  
31 62.5 17.3 419 2 S73846  
32 62 17.1 295 2 D42452  
33 62 17.1 1502 1 RGRH1  
34 61.5 16.9 481 2 S73846  
35 61.5 16.9 492 2 T44922  
36 61 16.9 587 2 J01419  
37 60.5 16.7 136 2 T2240  
38 60 16.7 201 2 A31380  
39 60 16.6 384 1 W24141  
40 60 16.6 447 2 S2437  
41 59 16.3 329 2 S73846  
42 59 16.3 329 2 S73846  
43 59 16.3 247 2 S26733  
44 58.5 16.2 266 2 T35449  
45 58.5 16.2 299 2 G64541

probable NADH oxid  
MG223 homolog F10  
Cl protein - tobac  
Cvcl/CYP3 transcri  
probable phosphos  
Cvcl/CYP3 transcri  
PC gamma (190) rec  
hypothetical prote  
probable flagellar  
E2 protein - human  
Cup-diacylglycerol  
5'-aminolevulinate  
hypothetical prote  
hypothetical prote  
cell binding facto

ALIGNMENTS

RESULT 1  
QCVCVL  
All protein - tomato golden mosaic virus  
C:Species: tomato golden mosaic virus  
A:Note: host Nicotiana sp. (tobacco)  
C:Date: 28-Aug-1985 #sequence\_revision 28-Aug-1985 #text\_change 08-Apr-1994  
EMBL J 3 2197-2205 1984  
E:Contig: A04170 Stejn, V.E.; Coutts, R.H.A.; Buck, K.W.  
A:Title: Complete nucleotide sequence of the infectious cloned DNA components of toma  
A:Reference number: A04163  
A:Accession: A04170  
A:Molecule type: DNA  
A:Residues: 1-352 <HAM>  
C:Comment: The genome consists of two circular, single-stranded DNA components, DNA A  
C:Genetics:  
A:Map position: segment A  
C:Superfamily: tomato golden mosaic virus All protein

Query Match 95.4% Score 349 DB 1 Length 352  
Best Local Similarity 95.7% Fixed NO. 1-30  
Matches 67 Conservative 1 Mismatches 2 Indels 0 Gaps 0

Qy 1 TLVGFQVQDGAAGAGCQTSDAAALNASSKEALQIIREKIPKYLQFQFHNLNSL 60  
Db 111 TLVGFQVQDGRSGGCQTSDAAALNASSKEALQIIREKIPKYLQFQFHNLNSL 170  
Qy 61 DRIFDKTEP 70  
Db 171 DRIFDKTEP 180

RESULT 2  
QCVCPT  
All protein - potato yellow mosaic virus (isolate Venezuela)  
C:Species: potato yellow mosaic virus  
C:Date: Jun-92 #sequence\_revision 30-Jun-1992 #text\_change 16-Jun-2000  
C:Accession: JH0364  
R:Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.  
J. Gen. Virol. 72, 1515-1520, 1991  
A:Title: The nucleotide sequence of the infectious cloned DNA components of potato ye  
A:Reference number: JH0362; MUID:91311403  
A:Accession: JH0364  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 361  
C:Genetics:  
A:Cross-references: GB:D00940; NID:g222459; PIDN:BA00782.1; PID:g222459  
A:Map position: segment A  
C:Superfamily: tomato golden mosaic virus All protein

```

Query Match      68.8%; Score 249; DB 1; Length 361;
Best Local Similarity 66.7%; Pred. No. 1e-19;
Matches 46; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

QY 1 TLVWGFQVDGAAAGGCGTSDNDAAEALNASSKEALQITREKIPKYLQFPHNLNSL 60
   |||||
Db 110 TVWGFQVDGSRGSGQGTVDNDAAEALNASSKEALQITREKIPKYLQFPHNLNSL 169
   |||||

QY 61 DRIFOKTPEP 69
   |||||
Db 170 DRIFOKAPE 178
   |||||

RESULT 3
S07594
hypothetical protein, 40.4k - cassava latent virus (Nigerian isolate)
C:Species: cassava latent virus
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
R:Morris, R.P., Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.
Nucleic Acids Res. 18, 197-198, 1990
A:Title: Nucleotide sequence of the infectious cloned DNA components of African cassava
A:Reference number: S07590; MUID:90174930
A:Accession: S07594
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 358 <NOPS>
A:Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA14953.1; PID:g59376
C:Genetics:
A:Map position: segment DNAL
C:Superfamily: tomato golden mosaic virus ALL protein

Query Match      65.2%; Score 236; DB 2; Length 358;
Best Local Similarity 60.0%; Pred. No. 2.6e-18;
Matches 42; Conservative 14; Mismatches 14; Indels 0; Gaps 0;

QY 1 TLVWGFQVDGAAAGGCGTSDNDAAEALNASSKEALQITREKIPKYLQFPHNLNSL 60
   |||||
Db 109 TVWGFQVDGSRGSGQGTSDNDAAEALNASSKEALQITREKIPKYLQFPHNLNSL 168
   |||||

QY 61 DRIFOKTPEP 70
   |||||
Db 169 DRIFOKPAP 178
   |||||

RESULT 4
JQ1887
ALL protein - tomato yellow leaf curl virus (strain Australia)
C:Species: tomato yellow leaf curl virus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C:Accession: JQ1887
R:Dr, I.B.; Rigen, J.B.; Krake, L.R.; Mullineux, P.M.; Rezaian, M.A.
J. Gen. Virol. 74, 147-151, 1993
A:Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.
A:Reference number: JQ1887; MUID:95139776
A:Accession: JQ1887
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-362 <DR>
A:Cross-references: GB:S53251
C:Superfamily: tomato golden mosaic virus ALL protein

Query Match      63.3%; Score 229; DB 1; Length 362;
Best Local Similarity 52.9%; Pred. No. 1.5e-17;
Matches 45; Conservative 10; Mismatches 14; Indels 16; Gaps 1;

QY 1 TLVWGFQVDGAAAGGCGTSDNDAAEALNASSKEALQITREKIPKYLQFPHNLNSL 60
   |||||

```

```

Db 110 TLWGFQVDGSRGSGQGTSDNDAAEALNASSKEALQITREKIPKYLQFPHNLNSL 169
   |||||
QY 61 DRI-----FDXTPE 69
   |||||
Db 170 DRITPPLEVYSPFLSSFDNPE 194
   |||||

RESULT 5
S39211
gene C1 protein - tomato yellow leaf curl virus
C:Species: tomato yellow leaf curl virus
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
R:Nakatsi, P.; Hidalgo, E.; Accotto, G.; Moriones, E.
submitted to the EMBL Data Library, August 1993
A:Description: High similarity among the tomato yellow leaf curl virus isolates from
A:Reference number: S39209
A:Accession: S39211
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 359 <NOPS>
A:Cross-references: EMBL:Y25751; NID:g433655; PIDN:CAA81026.1; PID:g433658
C:Superfamily: tomato golden mosaic virus ALL protein

Query Match      61.3%; Score 222; DB 2; Length 359;
Best Local Similarity 58.0%; Pred. No. 8.9e-17;
Matches 40; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

QY 2 LVWGFQVDGAAAGGCGTSDNDAAEALNASSKEALQITREKIPKYLQFPHNLNSL 61
   |||||
Db 111 LENGFTVIGDSRGSGQGTANDAYAKAINAGSKSEALDVIRELAPROYLURPHNINSNLD 170
   |||||

QY 62 RIFDXTPPEP 70
   |||||
Db 171 RVFQVPAP 179
   |||||

RESULT 6
JQ2300
replicase - pepper huasteco virus (component A)
N:Alternate names: ORF ALL protein
C:Species: pepper huasteco virus
C:Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Sep-1999
C:Accession: JQ2300
R:Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Bastrelia, L.; Rivera-Bustamante,
J. Gen. Virol. 74, 2225-2231, 1993
A:Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b
A:Reference number: JQ2299; MUID:94015007
A:Accession: JQ2300
A:Molecule type: DNA
A:Residues: 359 <NOPS>
A:Cross-references: GB:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
C:Superfamily: tomato golden mosaic virus ALL protein

Query Match      60.2%; Score 218; DB 2; Length 349;
Best Local Similarity 57.1%; Pred. No. 2.4e-16;
Matches 40; Conservative 14; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLVWGFQVDGAAAGGCGTSDNDAAEALNASSKEALQITREKIPKYLQFPHNLNSL 60
   |||||
Db 110 TVWGFQVDGSRGSGQGTANDAYAKAINAGSKSEALQITREKIPKYLQFPHNLNSL 169
   |||||

QY 61 DRIFOKTPEP 70
   |||||
Db 170 WRITQTPPEP 179
   |||||

RESULT 7
S31875
ALL protein - pepper rizado amarillo virus
C:Species: pepper rizado amarillo virus

```



A:Residues: 1-351 <HOM>  
A:Cross-references: EMBL:224758; NID:q93551; PID:CAA80B91.1; PID:q584046  
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 58.6%; Score 212; DB 2; Length 351;  
Best Local Similarity 39.7%; Pred No. 1,1e-15;  
Matches 40; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

QY 4 WGEFQVQAAAAAGCQTSNDAAAEALNASKSEFALQITREKIPKYLQFQHNLSNLDI 63  
Db 113 WGFQIQDGRSARGQSQSANDYAAALNASKSEALKILRLAPROYLRDHFHTSSNLDRI 172  
QY 64 FDKTPTTP 70  
Db 173 FTKPTPTTP 179

RESULT 10  
OCQW1

AVI protein - abutilon mosaic virus (isolate West India)  
C:Species: abutilon mosaic virus  
C:Accession: JQ1870 #sequence\_revision 31-Mar-1992 #text\_change 08-Apr-1994  
R:Frischmuth, T.; Zilmatt, G.; Jeske, H.  
Virolgy 178, 461-468, 1990  
A:Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as  
A:Reference number: A36214; MUID:91020984  
A:Accession: A36214  
A:Molecule type: DNA  
A:Residues: 1-355 <PRO>  
A:Cross-references: EMBL:X15983  
C:Genetics:  
A:Map position: segment A  
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 58.6%; Score 212; DB 1; Length 355;  
Best Local Similarity 52.9%; Pred No. 1,1e-15;  
Matches 39; Conservative 13; Mismatches 18; Indels 0; Gaps 0;

QY 1 TLVKEFGVGAAGAAGCQTSNDAAAEALNASKSEALQITREKIPKYLQFQHNLSNLI 60  
Db 110 TAMEGFEQIDGRSARGQSQFANQYAKLNAGDVSALNLIKKEQPKDYLVNHNIRSL 169  
QY 61 DFKETKPTTP 70  
Db 170 ERIFAKAPEP 179

RESULT 11  
JQ1870

AVI protein - tomato mottle virus (isolate Florida)  
C:Species: tomato mottle virus  
C:Accession: JQ1870 #sequence\_revision 17-Feb-1994 #text\_change 07-May-1999  
R:Abouzid, A.M.; Polston, J.E.; Hiebert, E.  
J. Gen. Virol. 73, 3225-3229, 1992  
A:Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated  
A:Reference number: JQ1869; MUID:93107958  
A:Accession: JQ1870  
A:Status: translation not shown  
A:Cross-references: EMBL:224758 <HOM>  
A:Residues: 1-358 <ARO>  
C:Genetics:  
A:Map position: segment A  
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 57.7%; Score 209; DB 1; Length 358;  
Best Local Similarity 52.9%; Pred No. 2,3e-15;  
Matches 37; Conservative 17; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLWGCFVQDGAAGCGCOTSDNAAAFALNASSKEEALQIIREKIPKYLQFPHLNSNL 60  
 DB 107 TIEKGFQIDGRSAGCGQTSKALNAGSVSALVLRKEDTNDYQIUNIRSNL 166  
 QY 61 DRIFQKTPPE 70  
 DB 167 BRIFAKAPEP 176  
 RESULT 12  
 S39235  
 A:Accession: S39235  
 A:Status: preliminary  
 A:Title: gene C1 protein - tomato yellow leaf curl virus  
 C:Species: tomato yellow leaf curl virus  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 08-Sep-1997  
 C:Accession: S39235  
 R:Crepi, S.; Norris, A.; Vaira, A.; Bosco, D.; Accotto, G.  
 submitted to the EMBL Data Library, December 1993  
 A:Reference: S39233  
 A:Accession: S39235  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-359 <R>  
 A:Cross-references: EMBL:Z28390; NID:g1041671; PID:g1334964  
 C:Superfamily: tomato golden mosaic virus AL1 protein  
 Query Match 57.7%; Score 209; DB 2; Length 359;  
 Best Local Similarity 53.6%; Pred. No. 2.4e-15;  
 Matches 37; Conservative 13; Mismatches 19; Indels 0; Gaps 0;  
 QY 2 LWGCFVQDGAAGCGCOTSDNAAAFALNASSKEEALQIIREKIPKYLQFPHLNSNL 61  
 DB 111 LKGFQIDGRSAGCGQTSKALNAGSVSALVLRKEDTNDYQIUNIRSNL 170  
 QY 62 RIFDKTPE 70  
 DB 171 KVFQVPPAP 179  
 RESULT 13  
 S38360  
 A:Accession: S38360  
 A:Status: preliminary  
 A:Title: AL1 protein - beet curly top virus  
 C:Species: beet curly top virus  
 C:Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 20-Sep-1999  
 C:Accession: S28360  
 R:Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.  
 EMO J. 5, 1761-1767, 1986  
 submitted to the EMBL Data Library, December 1993  
 A:Reference: S28360  
 A:Accession: S28360  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-385 <R>  
 A:Cross-references: GB:M24597; EMBL:X04144; NID:g210676; PID:AA42751.1; PID:g210679  
 C:Superfamily: tomato golden mosaic virus AL1 protein  
 Query Match 57.7%; Score 209; DB 2; Length 385;  
 Best Local Similarity 52.9%; Pred. No. 2.5e-15;  
 Matches 37; Conservative 17; Mismatches 16; Indels 0; Gaps 0;  
 QY 1 TLWGCFVQDGAAGCGCOTSDNAAAFALNASSKEEALQIIREKIPKYLQFPHLNSNL 60  
 DB 137 TIEKGFQIDGRSAGCGQTSKALNAGSVSALVLRKEDTNDYQIUNIRSNL 196  
 QY 61 DRIFQKTPPE 70  
 DB 197 OKIFQRPDP 206  
 RESULT 14  
 S39235

S59885  
 C:Species: tomato yellow leaf curl virus  
 C:Date: 14-Jan-1995 #sequence\_revision 01-Mar-1996 #text\_change 20-Sep-1999  
 C:Accession: S59885  
 R:Hong, Y.; Harrison, B.D.  
 submitted to the EMBL Data Library, February 1995  
 A:Description: Nucleotide sequences from tomato leaf curl viruses from different coun  
 d geminiviruses.  
 A:Reference number: S58346  
 A:Accession: S59885  
 A:Status: preliminary  
 A:Title: gene C1 protein - tomato yellow leaf curl virus  
 C:Species: tomato yellow leaf curl virus  
 C:Date: 06-Jan-1995 #sequence\_revision 01-Mar-1996 #text\_change 20-Sep-1999  
 C:Accession: S59885  
 A:Status: preliminary  
 A:Residues: 1-360 <R>  
 A:Cross-references: EMBL:Z48182; NID:g944838; PID:CAA88229.1; PID:g974211  
 C:Superfamily: tomato golden mosaic virus AL1 protein  
 Query Match 55.8%; Score 202; DB 2; Length 360;  
 Best Local Similarity 57.6%; Pred. No. 1.4e-14;  
 Matches 38; Conservative 13; Mismatches 15; Indels 0; Gaps 0;  
 QY 4 WGEFVQDGAAGCGCOTSDNAAAFALNASSKEEALQIIREKIPKYLQFPHLNSNL 63  
 DB 113 PGVQIDGRSAGCGQTSKALNAGSVSALVLRKEDTNDYQIUNIRSNL 172  
 QY 64 PDKTPE 69  
 DB 173 PTPSAE 178  
 RESULT 15  
 Q0CVCI  
 A:Accession: Q0CVCI  
 A:Status: preliminary  
 A:Title: AL1 protein - tomato yellow leaf curl virus  
 C:Species: tomato yellow leaf curl virus  
 C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 23-Jul-1999  
 C:Accession: D40779  
 R:Navot, N.; Pichersky, E.; Zeldin, M.; Zamir, D.; Czosnek, H.  
 Virology 185, 151-161, 1991  
 A:Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin  
 A:Reference number: A40779; MUID:92024070  
 A:Accession: D40779  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-357 <R>  
 A:Cross-references: GB:X15656; NID:g62204; PID:CAA33688.1; PID:g62207  
 C:Superfamily: tomato golden mosaic virus AL1 protein  
 Query Match 55.2%; Score 200; DB 1; Length 357;  
 Best Local Similarity 62.3%; Pred. No. 2.3e-14;  
 Matches 38; Conservative 11; Mismatches 12; Indels 0; Gaps 0;  
 QY 4 WGEFVQDGAAGCGCOTSDNAAAFALNASSKEEALQIIREKIPKYLQFPHLNSNL 63  
 DB 111 PGVQIDGRSAGCGQTSKALNAGSVSALVLRKEDTNDYQIUNIRSNL 170  
 QY 64 P 64  
 DB 171 F 171

Search completed: February 3, 2001, 02:17:37  
 Job time: 2205 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:19:38 ; Search time 109.09 Seconds  
(without alignments)  
11.523 Million cell updates/sec

Title: US-09-289-346a-2

Perfect score: 362

Sequence: 1 TLVWGFEQVDGAAGSQCT.....QFHNLSNLDIFDKTEP 70

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

1: /cgn2\_6/ptodata/1/aa/5A\_COMB.pep.\*

2: /cgn2\_6/ptodata/1/aa/5B\_COMB.pep.\*

3: /cgn2\_6/ptodata/1/aa/6\_COMB.pep.\*

4: /cgn2\_6/ptodata/1/aa/PCUS\_COMB.pep.\*

5: /cgn2\_6/ptodata/1/aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Score	Query	Length	ID	Description
1	215	59.4	359	3	US-08-809-103B-2
2	215	59.4	359	3	US-08-809-103B-4
3	215	59.4	359	3	US-08-809-103B-6
4	215	59.4	359	3	US-08-809-103B-8
5	68.5	18.9	1713	3	US-08-600-982-24
6	68.5	18.9	2161	4	US-08-455-543A-24
7	68.5	18.9	2161	4	US-08-455-543A-29
8	65.5	18.1	2161	1	US-08-455-543A-49
9	65.5	18.1	2161	1	US-08-455-543A-51
10	65.5	18.1	2161	2	US-08-223-305C-49
11	65.5	18.1	2161	2	US-08-223-305C-51
12	65.5	18.1	2161	2	US-08-311-363-2
13	63.5	17.5	334	5	5290690-11
14	59.5	16.3	454	3	US-08-446-100-11
15	59.5	16.3	454	3	US-08-446-100-31
16	58.5	16.2	439	1	US-07-637-870-9
17	58.5	16.2	439	1	US-07-637-399-6
18	58.5	16.2	439	1	US-08-112-703-6
19	58	16.0	844	3	US-09-029-267-20
20	57	15.7	454	3	US-08-446-100-30
21	56	15.5	454	3	US-08-446-100-25
22	56	15.5	454	3	US-08-446-100-26
23	56	15.5	454	3	US-08-446-100-28
24	56	15.5	454	3	US-08-446-100-29
25	56	15.5	548	2	US-08-871-266B-17
26	56	15.5	548	2	US-09-018-864A-17
27	56	15.5	548	3	US-08-871-267B-23
28	55.5	15.3	111	1	US-08-185-582-2

Query Match 59.4%; Score 215; DB 3; Length 359;  
Best Local Similarity 55.1%; Pred. No. 1.1e-19;  
Matches 38; Conservative 12; Mismatches 19; Indels 0; Gaps 0;

29 55.5 15.3 770 1 US-08-445-135-2  
30 54 14.9 485 1 US-07-881-075-1  
31 54 14.9 485 1 US-08-120-827-1  
32 54 14.9 485 1 US-08-478-675-1  
33 54 14.9 485 1 US-08-552-142A-13  
34 53.5 14.8 103 3 US-08-552-142A-13  
35 53.5 14.8 497 2 US-08-570-311-2  
36 53.5 14.8 457 2 US-08-353-485-2  
37 53.5 14.8 2628 2 US-08-570-311-14  
38 53 14.6 751 3 US-09-036-987A-24  
39 53 14.6 783 5 5231168-2  
40 52.5 14.5 268 2 US-08-897-980-186  
41 52.5 14.5 268 2 US-08-897-980-186  
42 52.5 14.5 268 3 US-09-095-855-186  
43 52.5 14.5 340 3 US-09-095-855-205  
44 52.5 14.5 439 1 US-07-762-681A-1  
45 52.5 14.5 921 1 US-08-188-582-2

#### ALIGNMENTS

RESULT 1  
US-08-809-103B-2  
: Sequence 2, Application US/08809103B  
: Patent No. 6133505  
: GENERAL INFORMATION:  
: INVENTOR: GREGORY B. BOUQU  
: TITLE OF INVENTION: ANTIGENIC DNA VIRUS RESISTANT  
: TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME  
: NUMBER OF SEQUENCES: 17  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: YOUNG & THOMPSON  
: STREET: 745 South 23rd Street  
: CITY: Arlington  
: STATE: Virginia  
: COUNTRY: U.S.A.  
: ZIP: 22202  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patent Release #1.0, Version #1.30  
: COUNTRY: U.S.A.  
: APPLICATION NUMBER: US/08/809, 103B  
: FILING DATE: 17-MAR-1997  
: CLASSIFICATION: 800  
: PRIOR APPLICATION DATA:  
: FILING DATE: 15-SEP-1994  
: PRIOR APPLICATION DATA: NO PCT/FR95/01192  
: FILING DATE: 15-SEP-1995  
: ATTORNEY/AGENT INFORMATION:  
: NAME: PATCH, Andrew J.  
: REGISTRATION NUMBER: 32,925  
: REFERENCE/POCKET NUMBER: US944L CNR TOM  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (703) 605-0529  
: TELEFAX: (703) 605-0573  
: TELEX: 248425 PMON  
: INFORMATION FOR SEQ ID NO: 2:  
: LENGTH: 359 amino acids  
: TYPE: amino acid  
: TOPOLOGY: linear  
: MOLECULAR TYPE: protein  
US-08-809-103B-2

QY 2 LVNGFQVGGAAAGCCTNDAAALNASSKEALQIREKIPKYLQFHNLSNLD 61  
 Db 111 LVNGFTQDGSAGCGQTANDAYAKAINAGSKQALDVKELAPRDYVLFHFNLSNLD 170  
 QY 62 RIFDKTPEP 70  
 Db 171 KVFQVPEAP 179

RESULT 2 103b-4  
 US-08-809-103b-4  
 : Sequence 6, Application US/08809103B  
 : Patent No. 6133505  
 : GENERAL INFORMATION:  
 : APPLICANT: GRONENBORN, Bruno  
 : TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT  
 : TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME  
 : NUMBER OF SEQUENCES: 17  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: YOUNG & THOMPSON  
 : STREET: 745 South 23rd Street  
 : CITY: Arlington  
 : STATE: Virginia  
 : COUNTRY: U.S.A.  
 : ZIP: 22202  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: PatentIn Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/809.103B  
 : FILING DATE: 17-MAR-1997  
 : CLASSIFICATION: 800  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: FR 94.11040  
 : FILING DATE: 15-SEP-1994  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: PATCH, Andrew J.  
 : REGISTRATION NUMBER: 32.925  
 : REFERENCE/DOCKET NUMBER: US94AL CNR TOM  
 : TELEPHONE: (703) 521-2297  
 : TELEX: 248425 EMBON  
 : INFORMATION FOR SEQ ID NO: 4:  
 : LENGTH: 359 amino acids  
 : TYPE: amino acid  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: protein  
 : US-08-809-103b-4

Query Match 59.4%; Score 215; DB 3; Length 359;  
 Best Local Similarity 55.1%; Pred. No. 1.le-19;  
 Matches 38; Conservative 12; Mismatches 19; Indels 0;  
 QY 2 LVNGFQVGGAAAGCCTNDAAALNASSKEALQIREKIPKYLQFHNLSNLD 61  
 Db 111 LVNGFTQDGSAGCGQTANDAYAKAINAGSKQALDVKELAPRDYVLFHFNLSNLD 170  
 QY 62 RIFDKTPEP 70  
 Db 171 KVFQVPEAP 179

RESULT 3

US-08-809-103b-6

: Sequence 6, Application US/08809103B  
 : Patent No. 6133505  
 : GENERAL INFORMATION:  
 : APPLICANT: GRONENBORN, Bruno  
 : TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT  
 : TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME  
 : NUMBER OF SEQUENCES: 17  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: YOUNG & THOMPSON  
 : STREET: 745 South 23rd Street  
 : CITY: Arlington  
 : STATE: Virginia  
 : COUNTRY: U.S.A.  
 : ZIP: 22202  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: PatentIn Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/809.103B  
 : FILING DATE: 17-MAR-1997  
 : CLASSIFICATION: 800  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: FR 94.11040  
 : FILING DATE: 15-SEP-1994  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: PATCH, Andrew J.  
 : REGISTRATION NUMBER: 32.925  
 : REFERENCE/DOCKET NUMBER: US94AL CNR TOM  
 : TELEPHONE: (703) 521-2297  
 : TELEX: 248425 EMBON  
 : INFORMATION FOR SEQ ID NO: 6:  
 : LENGTH: 359 amino acids  
 : TYPE: amino acid  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: protein  
 : US-08-809-103b-6

Query Match 59.4%; Score 215; DB 3; Length 359;  
 Best Local Similarity 55.1%; Pred. No. 1.le-19;  
 Matches 38; Conservative 12; Mismatches 19; Indels 0; Gaps 0;

QY 2 LVNGFQVGGAAAGCCTNDAAALNASSKEALQIREKIPKYLQFHNLSNLD 61  
 Db 111 LVNGFTQDGSAGCGQTANDAYAKAINAGSKQALDVKELAPRDYVLFHFNLSNLD 170  
 QY 62 RIFDKTPEP 70  
 Db 171 KVFQVPEAP 179

RESULT 4

: Sequence 8, Application US/08809103B  
 : Patent No. 6133505  
 : GENERAL INFORMATION:  
 : APPLICANT: GRONENBORN, Bruno  
 : TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT  
 : TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME  
 : NUMBER OF SEQUENCES: 17  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: YOUNG & THOMPSON  
 : STREET: 745 South 23rd Street  
 : CITY: Arlington  
 : STATE: Virginia

```

; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; FILE: FPIF0001.DAT
; COMPUTER: IBM PC
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/809,103B
; FILING DATE: 17-MAR-1997
; CLASSIFICATION: 800
; PRIORITY: 800
; APPLICATION NUMBER: FR 94.11040
; FILING DATE: 15-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR95/01192
; FILING DATE: 15-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: PATCH, Andrew J., 925
; ADDRESS: 10000 N. 11TH AVE.
; REFERENCE/DOCKET NUMBER: US94AL CNR TOM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 521-2297
; TELEFAX: (703) 685-0573
; TELEX: 248425 EMBON
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1713 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-809-103B-8

Query Match 59.4%; Score 215; DB 3; Length 359;
Local Similarity 50.1%; Pct 100; Indels 0;
Matches 38; Conservative 12; Mismatches 19; Gaps 0;

Oy 2 LWSRFQVDAAGAGCGQTSDNAAAEALNASSKEEALQIIREKIPERYLFQFINLNSNLD 61
Db 111 LLEWTFQDGRSGARGQQTANDATAYAKAINAGSKSQALQVKEIAPRDYVYLFHFNINSNLD 170

Oy 62 RLEKDTPEP 70
Db 171 KVFQWPEAP 179

RESULT 5
US-08-600-982-24
; Sequence 24, Application US/08600982
; GENERAL INFORMATION:
; ATTORNEY/AGENT INFORMATION:
; APPLICANT: Carter, William G.
; APPLICANT: Gil, Susanna A.
; TITLE OF INVENTION: Epiligrin, an Epithelial Ligand for
; TITLE OF INVENTION: Epiligrin
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESS: Johnson, O'Connor, Johnson, and Kindness
; STREET: 1420 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101-8100
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/10261A
; FILING DATE: 02-SEP-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Johnson, O'Connor, Johnson, and Kindness
; REGISTRATION NUMBER: 26,997
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 682-8100
; TELEFAX: (206) 224-0779
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1713 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: E170 protein as translated from sequence of
; DESCRIPTION: FIGURES 15A-15F, and as shown also in FIGURES 19A-19F
; FILING DATE: 02-SEP-1994

```

```

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Shelton, Dennis K.
; REGISTRATION NUMBER: 26,997
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 682-8100
; TELEFAX: (206) 224-0779
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1713 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: E170 protein as translated from sequence
; DESCRIPTION: of FIGURES 15A-15F, and as shown also in FIGURES
; DESCRIPTION: 19A-19F
; US-08-600-982-24

Query Match 18.9%; Score 68.5; DB 3; Length 1713;
Local Similarity 36.1%; Pct 100; Indels 7; Gaps 3;
Matches 27; Conservative 11; Mismatches 16;

Oy 9 VDAATATAYENILNIAKAAEDAAANRAASASALQTVIKEDLPK----AKTLSSNDSKLLN 608
Db 553 VDAATATAYENILNIAKAAEDAAANRAASASALQTVIKEDLPK----AKTLSSNDSKLLN 608

Oy 66 K 66
Db 609 K 609

RESULT 6
PCT-US94-10261A-24
; Sequence 24, Application PCT/US9410261A
; GENERAL INFORMATION:
; ATTORNEY/AGENT INFORMATION:
; APPLICANT: Carter, William G.
; APPLICANT: Gil, Susanna A.
; TITLE OF INVENTION: Epiligrin, an Epithelial Ligand for
; TITLE OF INVENTION: Epiligrin
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESS: Johnson, O'Connor, Johnson, and Kindness
; STREET: 1420 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101-8100
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/10261A
; FILING DATE: 02-SEP-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Johnson, O'Connor, Johnson, and Kindness
; REGISTRATION NUMBER: 26,997
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 682-8100
; TELEFAX: (206) 224-0779
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1713 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: E170 protein as translated from sequence of
; DESCRIPTION: FIGURES 15A-15F, and as shown also in FIGURES 19A-19F
; FILING DATE: 02-SEP-1994

```

Query Match 18.1%; Score 65.5; DB 1; Length 2161;  
Best Local Similarity 28.4%; Pred. No. 15;  
Matches 21; Conservative 5; Mismatches 19; Indels 29; Gaps 1;

RESULT 8  
US-08-455-543A-49

```

: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: FRAGMENT TYPE: internal
US-08-455-543A-51

Query Match
Best Local Similarity 28.4% Pred. No. 15:
Matches 21: Conservative 5: Mismatches 19: Indels 25: Gaps

QY 4 WEFQVDCAARAAGCCTCS-----NDAAAEALNASK 34
DB 707 WNAWYDGMAYGSSGSMVCYIFILFTCGNTILLAVFLAVDNLDAESLNTAOK 766
OY 35 EEALQITREKTPK 48
DY 767 EEAEKERRKKIARK 780

RESULT 10
US-08-223-305C-49
: Accession: 49 Application US/08223305C
: Patient No. 5651824
: GENERAL INFORMATION:
: APPLICANT: Harpold, Michael
: APPLICANT: Ellis, Steven
: APPLICANT: Williams, Mark
: APPLICANT: Feldman, Daniel
: APPLICANT: McCue, Ann
: APPLICANT: Harpold, Robert
: TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
: TITLE OF INVENTION: METHODS
: NUMBER OF SEQUENCES: 57
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Brown, Martin, Haller & McClain
: STREET: 1660 Union Street
: CITY: San Diego
: STATE: California
: COUNTRY: USA
: ZIP: 92101-2926
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ Version 1.5
: COMMENTS: CDS FROM DISKETTE
: APPLICATION NUMBER: US/08/223,305C
: FILING DATE: April 4, 1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/868,354
: FILING DATE: April 10, 1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/745,206
: FILING DATE: 15-JUG-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/620,250
: FILING DATE: 30-NOV-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/482,384
: FILING DATE: 20-FEB-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/603,751
: FILING DATE: 04-APR-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WO PCT/US89/01408
: FILING DATE: 04-APR-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/176,899
: FILING DATE: 04-APR-1988
: ATOMEX/SEQUENCE INFORMATION:
: NAME: Sequences
: REGISTRATION NUMBER: 33,779

```

REFERENCE/DOCKET NUMBER: 52516 (P519739)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619)238-0999  
FILING DATE: 04-APR-1989  
INFORMATION FOR SEQ ID NO: 2 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2161 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: Internal  
US-08-223-305C-49

Query Match 18.1%; Score 65.5; DB 2; Length 2161;  
Best Local Similarity 28.4%; Pred. No. 15;  
Matches 21; Conservative 5; Mismatches 19; Indels 29; Gaps 1;  
OY 4 WGEFVGCAARAGCCOTS-----NDAAAFALNASSK 34  
DB 707 NNAVMYDGMVGGSSGMVICVFIILFCGNYILLAVDLAVDLADAESLNTAQK 766

OY 35 PEALQIIREKPEK 48  
DB 767 EBAEKERKKIARK 780

RESULT 11  
US-08-223-305C-51  
Sequence 51, Application US/08223305C  
Patent No. 5851824  
GENERAL INFORMATION:  
APPLICANT: Harpold, Michael  
INVENTOR: Stephen L. Seidman  
APPLICANT: Williams, Mark  
APPLICANT: Feldman, Daniel  
APPLICANT: McCue, Ann  
APPLICANT: Brenner, Robert  
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND  
METHODS  
NUMBER OF SEQUENCES: 57  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Brown, Martin, Haller & McClain  
STREET: 1660 Union Street  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-2926  
COMPUTER: IBM Compatible  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/223,305C  
FILING DATE: April 4, 1994  
PRIORITY DATE: April 10, 1992  
APPLICATION NUMBER: 07/868,354  
FILING DATE: April 10, 1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/745,206  
FILING DATE: 15-AUG-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/620,250  
FILING DATE: 30-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/482,384  
FILING DATE: 20-FEB-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/603,751  
FILING DATE: 04-APR-1989  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US89/01408  
FILING DATE: 04-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/176,899  
FILING DATE: 04-APR-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Seidman, Stephanie L.  
REGISTRATION NUMBER: 33,779  
REGISTRATION/DOCKET NUMBER: 52516 (P519739)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619)238-0999  
FILING DATE: 04-APR-1989  
INFORMATION FOR SEQ ID NO: 51:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2161 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: Internal  
US-08-223-305C-51

Query Match 18.1%; Score 65.5; DB 2; Length 2161;  
Best Local Similarity 28.4%; Pred. No. 15;  
Matches 21; Conservative 5; Mismatches 19; Indels 29; Gaps 1;  
OY 4 WGEFVGCAARAGCCOTS-----NDAAAFALNASSK 34  
DB 707 NNAVMYDGMVGGSSGMVICVFIILFCGNYILLAVDLAVDLADAESLNTAQK 766

OY 35 PEALQIIREKPEK 48  
DB 767 EBAEKERKKIARK 780

RESULT 12  
US-08-311-363-2  
Sequence 2, Application US/08311363  
Patent No. 5876958  
GENERAL INFORMATION:  
APPLICANT: Harpold, Michael  
INVENTOR: Stephen L. Seidman  
APPLICANT: Williams, Mark  
APPLICANT: Feldman, Daniel  
APPLICANT: McCue, Ann  
APPLICANT: Brenner, Robert  
TITLE OF INVENTION: Human Calcium Channel Compositions and  
Methods  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Brown, Martin, Haller & McClain  
STREET: 1660 Union Street  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-2926  
COMPUTER: IBM PC compatible  
MEDIUM TYPE: 5.25 disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/311,363  
FILING DATE: 15-AUG-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/745,206  
FILING DATE: 15-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Seidman, Stephanie L.  
REGISTRATION NUMBER: 33,779  
REGISTRATION/DOCKET NUMBER: 6362-51506  
TELECOMMUNICATION INFORMATION:



```

; TELEPHONE: (619)238-0999
; TELEFAX: (619)238-0062
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE LENGTH: 2161
; LENGTH: 2161 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
us-08-311-363-2

Query Match 18.18: Score 65.5; DB 2; Length 2161:
Best Local Similarity 28.4%; Pred. No. 15;
Matches 21; Conservative 5; Mismatches 19; Indels 29; Gaps 1;

QY 4 WGEFQVDGAAGAGCCTTS-----NDAAALNASSK 34
DB 707 WNAVYDGMATYGGPSSGMIVCIYFIILFCGNYILLNVFLAVDLAIDAESLNTAOK 766
QY 35 BENQIIREKPEK 48
DB 767 BEASEKERKAKRK 780

RESULT 13
5290690-11 5290690
; PATENT APPLICANT: MABET, NADIR; LASTERS, IGNACE; STANSSENS, PATRICK
; MATTHYSSENS, GASTON; MODAK, SHOSHANA; QUAX, WILHELMUS J.
; TITLE OF INVENTION: METHODS AND MEANS FOR CONTROLLING THE
; STABILITY OF PROTEINS
; NUMBER OF SEQUENCES: 22
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/398,706
; FILING DATE: 25-AUG-1989
; SEQ ID NO: 11
; LENGTH: 334
5290690-11

Query Match 17.5%; Score 63.5; DB 5; Length 334:
Best Local Similarity 28.4%; Pred. No. 2.4; 27; Indels 9; Gaps 1;
Matches 19; Conservative 12; Mismatches 27; Indels 9; Gaps 1;

QY 2 LVNKEFQVDGAAGAGCCTTSNDAAALNASSK-----ENQIIREKIPKYLQ 52
DB 82 LAMEIGLDIVVSTGRTKREDAAKHLEAGAKKVIISAPAKVENITVYVNGQKIDPK 141
QY 53 FHLNLSN 59
DB 142 AHRVSN 148

RESULT 14
us-08-672-814D-11
; Sequence 11, Application US/08672814D
; Patent No. 5552480
; GENERAL INFORMATION:
; APPLICANT: David W.
; APPLICANT: Mark Y. Peltia
; TITLE OF INVENTION: MAMMALIAN
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cell Therapeutics, Inc.
; STREET: 201 Elliott Avenue West
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98119
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" disk, 1.44Mb, double side, high density
; COMPUTER: PC Clone (486 microprocessor)
; OPERATING SYSTEM: MS-DOS Version 6.1, Windows NT

; SOFTWARE: WORD 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,814D
; FILING DATE: 28-JUN-1996
; ANNOTATION/PATENT INFORMATION:
; NAME: Paciszewski, Stephen
; REGISTRATION NUMBER: 36,131
; REFERENCE/DOCKET NUMBER: 1803
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)284-6206
; TELEFAX: (206)284-6206
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE LENGTH: 446
; STRONG CHARACTERISTICS:
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: no
; NATURALNESS: no
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM: Drosophila
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; GROSS TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; FEATURE:
; NAME/KEY: CDP-diacylglycerol synthase (CDS)
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
us-08-672-814D-11

Query Match 16.6%; Score 60; DB 2; Length 446:
Best Local Similarity 34.0%; Pred. No. 9.9;
Matches 17; Conservative 9; Mismatches 22; Indels 2; Gaps 1;

QY 20 TSDAAALNASSKEENQIIREKIPKYLQFHLNLSNLRIDPEKTP 69
DB 20 SGSDAANKRNAASDSDHVDSEEEKIPKEKVD--SLAKNLQGGTKDTP 67

RESULT 15
us-08-446-100-31
; Sequence 31, Application US/08446100
; Patent No. 5606525
; GENERAL INFORMATION:
; APPLICANT: Broomfield, Clarence A
; APPLICANT: Millard, Charles B
; APPLICANT: Lockridge, Oksana
; TITLE OF INVENTION: Site-Directed Mutagenesis of Esterases
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Broomfield, Clarence and Assoc.
; STREET: 9669 A Main Street, P.O. Box 2509
; CITY: Fairfax
; STATE: VA
; COUNTRY: US
; ZIP: 22031
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,100
; FILING DATE: 19-MAY-1995
```

```

:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
:   NAME: Hendricks, Glenna
:   REFERENCE NUMBER: 32,535
:   REFERENCE/DOCKET NUMBER: ocmfield
: TELECOMMUNICATION INFORMATION:
:   TELEPHONE: (703) 425-4250
:   TELEFAX: (703) 425-2767
: INFORMATION FOR SEQ ID NO: 31:
:   SEQUENCE CHARACTERISTICS:
:     LENGTH: 454 amino acids
:     TYPE: amino acid
:     STRANDEDNESS: single
:     TOPOLOGY: unknown
:   MOLECULE TYPE: protein
:   HYPOTHETICAL: YES
:   ANTI-SENSE: YES
:   FRAGMENT TYPE: N-terminal
:   ORIGINAL SOURCE:
:     ORGANISM: human esterases
:   US-08-446-100-31

```

```

Query Match      16.3%; Score 59; DB 3; Length 454;
Best Local Similarity 31.8%; Pred. No. 14;
Matches 21; Conservative 10; Mismatches 25; Indels 10; Gaps 4;

QY  13 AAAGCGTSDNDAA-AEAINNASKKEALQIIREKIPKYLQF-----INLNSLD-HI 63
    |||:  |||:  |||:  |||:  |||:  |||:  |||:  |||:  |||:  |||:
Db  156 AITAGCKTTTSAANVHCLQKTEELLETTL-KIGNSYLWYRETQRESHLJGTVIDGML 214
    |||:  |||:  |||:  |||:  |||:  |||:  |||:  |||:  |||:  |||:
QY  64 FDKTPE 69
    ||||
Db  215 LKTPPE 220

```

Search completed: February 3, 2001, 02:19:40  
 Job time: 1702 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:15:25 ; Search time 144.12 Seconds  
(without alignments)  
16,608 Million cell updates/sec

Title: US-09-289-346a-2  
Protein score: 266485  
Sequence: 1 TLWKEQVNDGAAGCQT.....PQHNLNSLDRFDKTPETP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues  
Listing first 45 summaries

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Database : A: Geneseq.36.\*  
1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.\*

## SUMMARIES

Result	Score	Query	Length	ID	Description
1	215	59.4	353	18	Bean golden mosaic
2	215	59.4	353	18	Bean golden mosaic
3	215	59.4	353	18	Bean golden mosaic
4	215	59.4	353	18	Bean golden mosaic
5	215	59.4	353	18	Bean golden mosaic
6	215	59.4	353	17	Sardinian tomato y
7	215	59.4	359	17	Sardinian tomato y
8	215	59.4	359	17	ORF 4 gene product
9	213	58.8	353	8	Tomato mottle virus
10	213	58.8	361	18	Tomato mottle virus
11	213	58.8	361	18	Tomato mottle virus
12	213	58.8	361	18	Tomato mottle virus

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

RESULT	ID	W34338 standard: Protein: 353 AA.	13	213	58.8	361	18	W34326
W34338	XX		14	206.5	57.0	361	8	P70562
XX	XX		15	201	55.5	362	19	W56495
XX	XX		16	200	55.2	357	18	W34329
XX	XX		17	200	55.2	357	18	W34330
XX	XX		18	200	55.2	357	18	W34331
XX	XX		19	192	53.0	357	18	W34337
XX	XX		20	69.5	19.2	512	19	W68473
XX	XX		21	68.5	18.9	1713	16	R70148
XX	XX		22	65.5	18.1	619	13	R27651
XX	XX		23	65.5	18.1	2161	14	R33545
XX	XX		24	65.5	18.1	2161	16	R71001
XX	XX		25	65.5	18.1	2161	19	W63197
XX	XX		26	65.5	18.1	2161	19	W63197
XX	XX		27	65.5	18.0	292	21	Y92319
XX	XX		28	65	18.0	335	21	Y92318
XX	XX		29	65	18.0	335	21	Y92318
XX	XX		30	64.5	17.8	299	18	W55450
XX	XX		31	63.5	17.5	334	11	R03211
XX	XX		32	60.5	16.7	111	20	X37827
XX	XX		33	60.5	16.7	299	20	W89829
XX	XX		34	60.5	16.7	299	20	W89829
XX	XX		35	60.5	16.7	299	20	W89849
XX	XX		36	59.5	16.4	673	21	Y58996
XX	XX		37	59	16.3	454	21	Y49470
XX	XX		38	58.5	16.2	299	19	W98322
XX	XX		39	58.5	16.2	299	19	W73034
XX	XX		40	57.8	15.0	944	18	W09773
XX	XX		41	57.3	15.0	944	18	W09773
XX	XX		42	57.3	15.9	266	20	Y21957
XX	XX		43	57	15.7	454	21	Y49500
XX	XX		44	56.5	15.6	249	20	W89822
XX	XX		45	56.5	15.6	269	20	W89860

ALIGNMENTS

Bean golden mosaic geminivirus C1 protein.

Geminivirus; BCMV; C1 gene; transdominant mutation; transgenic plant; disease resistance.

Bean golden mosaic virus type II isolate Guatemala.

W09739110-AL.

23-OCT-1997.

15-APR-1997; 97WO-US05300.

16-APR-1996; 96US-0015517.

(SEMI-) SEMINIS VEGETABLE SEEDS INC. (WISC) WISCONSIN ALUMNI RES FOUND.

Ahlquist PG, Hanson SF, Lou HT, Maxwell DP, Stout JT; WPT: 1997-528447/48.

N-PSDB: T93314.

Transgenic plants expressing geminivirus AC1 and C1 wild-type and mutant genes - have increased resistance to geminivirus infection e.g. tomato mottle virus, tomato yellow leaf curl virus or bean golden mosaic geminivirus



CC highly conserved DNA-nicking and/or the NTP-binding domains.

```

XX XX
SQ Sequence 353 AA:

Query Match
Best Local Similarity 59.4%; Score 215; DB 18; Length 353;
Matches 40; Conservative 14; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLVWGFEVDGAAAGCGQTSNDMAAEALNASSKEALQIREKIPEKYLQFHNLSML 60
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 110 TLWGFQVdgrsarggqgqandsyakaInadslesaltlIkkepdkyViqhnrsl 169
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 DRIFDKTPPEP 70
   :|||:|||||
Db 170 erifvkvpep 179
   :|||:|||||

RESULT 4
W34334
ID W34334 standard; Protein; 353 AA.
XX AC W34334;
XX XX 27-APR-1998 (first entry)
XX DE Bean golden mosaic geminivirus Cl BG4228 mutant.
XX KW Geminivirus; BGWV; Cl gene: transdomloant mutation;
XX KW transgenic plant; disease resistance.
XX OS Bean golden mosaic virus type II isolate Guatemala.
XX XX W09739110-A1.
XX PD 23-OCT-1997.
XX PF 15-APR-1997; 97WO-US06300.
XX PR 16-APR-1996; 96US-0015517.
XX XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.
XX PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX WPI; 1997-526447/48.
XX DR N-PSDB: T93292.
XX XX Transgenic plants expressing geminivirus AC1 and CI wild-type and
XX XX mutant genes. e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
XX XX golden mosaic geminivirus
XX XX Example 5; Page 111-112; 132pp; English.
XX CC This protein comprises a transdominant lethal mutant of the bean
XX CC golden mosaic virus (BGWV) Cl protein (see W34334) that is required
XX CC for replication. It is encoded by a mutated CI open reading frame
XX CC BG4C228 (see T93292) and carries a mutation in the NTP-binding
XX CC domain. The invention involves production of transgenic plants
XX CC containing DNA comprising geminivirus CI or AC1 wild-type or mutant
XX CC sequences that negatively interfere in trans with geminiviral
XX CC replication during infection. Such transgenic plants are resistant
XX CC to viral infection. The AC1/CI genes are especially from BGWV,
XX CC tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
XX CC and encode polypeptides (see W34324-35) that have mutations in the
XX CC highly conserved DNA-nicking and/or the NTP-binding domains.
XX SQ Sequence 353 AA;

Query Match
Best Local Similarity 59.4%; Score 215; DB 18; Length 353;
Matches 40; Conservative 14; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLVWGFEVDGAAAGCGQTSNDMAAEALNASSKEALQIREKIPEKYLQFHNLSML 60
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 110 TLWGFQVdgrsarggqgqandsyakaInadslesaltlIkkepdkyViqhnrsl 169
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 DRIFDKTPPEP 70
   :|||:|||||
Db 170 erifvkvpep 179
   :|||:|||||

RESULT 4
W34334
ID W34334 standard; Protein; 353 AA.
XX AC W34334;
XX XX 27-APR-1998 (first entry)
XX DE Bean golden mosaic geminivirus Cl BG4228 mutant.
XX KW Geminivirus; BGWV; Cl gene: transdomloant mutation;
XX KW transgenic plant; disease resistance.
XX OS Bean golden mosaic virus type II isolate Guatemala.
XX XX W09739110-A1.
XX PD 23-OCT-1997.
XX PF 15-APR-1997; 97WO-US06300.
XX PR 16-APR-1996; 96US-0015517.
XX XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.
XX PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX WPI; 1997-526447/48.
XX DR N-PSDB: T93292.
XX XX Transgenic plants expressing geminivirus AC1 and CI wild-type and
XX XX mutant genes. e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
XX XX golden mosaic geminivirus
XX XX Example 5; Page 111-112; 132pp; English.
XX CC This protein comprises a transdominant lethal mutant of the bean
XX CC golden mosaic virus (BGWV) Cl protein (see W34334) that is required
XX CC for replication. It is encoded by a mutated CI open reading frame
XX CC BG4C228 (see T93292) and carries a mutation in the NTP-binding
XX CC domain. The invention involves production of transgenic plants
XX CC containing DNA comprising geminivirus CI or AC1 wild-type or mutant
XX CC sequences that negatively interfere in trans with geminiviral
XX CC replication during infection. Such transgenic plants are resistant
XX CC to viral infection. The AC1/CI genes are especially from BGWV,
XX CC tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
XX CC and encode polypeptides (see W34324-35) that have mutations in the
XX CC highly conserved DNA-nicking and/or the NTP-binding domains.
XX SQ Sequence 353 AA;

```

```

Best Local Similarity 57.1%; Pred. No. 1.4e-18;
Matches 40; Conservative 14; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLVWGFEVDGAAAGCGQTSNDMAAEALNASSKEALQIREKIPEKYLQFHNLSML 60
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 110 TLWGFQVdgrsarggqgqandsyakaInadslesaltlIkkepdkyViqhnrsl 169
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 DRIFDKTPPEP 70
   :|||:|||||
Db 170 erifvkvpep 179
   :|||:|||||

RESULT 5
W34335
ID W34335 standard; Protein; 353 AA.
XX AC W34335;
XX XX 27-APR-1998 (first entry)
XX DE Bean golden mosaic geminivirus Cl BG4262 mutant.
XX KW Geminivirus; BGWV; Cl gene: transdominant mutation;
XX KW transgenic plant; disease resistance.
XX OS Bean golden mosaic virus type II isolate Guatemala.
XX XX W09739110-A1.
XX PD 23-OCT-1997.
XX PF 15-APR-1997; 97WO-US06300.
XX PR 16-APR-1996; 96US-0015517.
XX XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.
XX PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX WPI; 1997-526447/48.
XX DR N-PSDB: T93293.
XX XX Transgenic plants expressing geminivirus AC1 and CI wild-type and
XX XX mutant genes. e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
XX XX golden mosaic geminivirus
XX XX Example 5; Page 115-116; 132pp; English.
XX CC This protein comprises a transdominant lethal mutant of the bean
XX CC golden mosaic virus (BGWV) Cl protein (see W34335) that is required
XX CC for replication. It is encoded by a mutated CI open reading frame
XX CC BG4C262 (see T93293) and carries a mutation in the NTP-binding
XX CC domain. The invention involves production of transgenic plants
XX CC containing DNA comprising geminivirus CI or AC1 wild-type or mutant
XX CC sequences that negatively interfere in trans with geminiviral
XX CC replication during infection. Such transgenic plants are resistant
XX CC to viral infection. The AC1/CI genes are especially from BGWV,
XX CC tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
XX CC and encode polypeptides (see W34324-35) that have mutations in the
XX CC highly conserved DNA-nicking and/or the NTP-binding domains.
XX SQ Sequence 353 AA;

Query Match
Best Local Similarity 59.4%; Score 215; DB 18; Length 353;
Matches 40; Conservative 14; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLVWGFEVDGAAAGCGQTSNDMAAEALNASSKEALQIREKIPEKYLQFHNLSML 60
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 110 TLWGFQVdgrsarggqgqandsyakaInadslesaltlIkkepdkyViqhnrsl 169
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 DRIFDKTPPEP 70
   :|||:|||||
Db 170 erifvkvpep 179
   :|||:|||||

RESULT 5
W34335
ID W34335 standard; Protein; 353 AA.
XX AC W34335;
XX XX 27-APR-1998 (first entry)
XX DE Bean golden mosaic geminivirus Cl BG4262 mutant.
XX KW Geminivirus; BGWV; Cl gene: transdominant mutation;
XX KW transgenic plant; disease resistance.
XX OS Bean golden mosaic virus type II isolate Guatemala.
XX XX W09739110-A1.
XX PD 23-OCT-1997.
XX PF 15-APR-1997; 97WO-US06300.
XX PR 16-APR-1996; 96US-0015517.
XX XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.
XX PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX WPI; 1997-526447/48.
XX DR N-PSDB: T93293.
XX XX Transgenic plants expressing geminivirus AC1 and CI wild-type and
XX XX mutant genes. e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
XX XX golden mosaic geminivirus
XX XX Example 5; Page 115-116; 132pp; English.
XX CC This protein comprises a transdominant lethal mutant of the bean
XX CC golden mosaic virus (BGWV) Cl protein (see W34335) that is required
XX CC for replication. It is encoded by a mutated CI open reading frame
XX CC BG4C262 (see T93293) and carries a mutation in the NTP-binding
XX CC domain. The invention involves production of transgenic plants
XX CC containing DNA comprising geminivirus CI or AC1 wild-type or mutant
XX CC sequences that negatively interfere in trans with geminiviral
XX CC replication during infection. Such transgenic plants are resistant
XX CC to viral infection. The AC1/CI genes are especially from BGWV,
XX CC tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
XX CC and encode polypeptides (see W34324-35) that have mutations in the
XX CC highly conserved DNA-nicking and/or the NTP-binding domains.
XX SQ Sequence 353 AA;

```

Qy	61	DRIFDKTP	70
		..	
Db	170	erifkvpep	179
		.....	.....
RESULT	6		
RD	R8870	standard; Protein; 359 AA.	
XX	XX	AC	R8870;
XX	XX	AC	R8870;
DT	07-NOV-1996	(first entry)	
XX	XX		
DE	Sardinian tomato yellow leaf curl virus mutated Rep protein (K227A).		
KW	Nucleotide triphosphate binding site; DNA helicase; RNA helicase; modification; mutation; viral replication; deficient; inhibition;		
KW	transgenic Nicotiana benthamiana plants generated by an		
KW	Sardinian isolate; STVLCV; transgenic plant; P-loop; C1 protein;		
KW	AL1 protein; dominant negative phenotype.		
XX	XX		
OS	Sardinian tomato yellow leaf curl virus.		
XX	XX		
XX	Key	Location/Qualifiers	
FT	Misc-difference	227	
XX	XX	/note= "wild-type Lys has been replaced by Ala"	
PN	XX	WO9608573-A1.	
XX	XX		
PD	21-MAR-1996.		
XX	XX		
PF	15-SEP-1995;	95WO-FR01192.	
XX	XX		
PF	15-SEP-1994;	94FR-0011040.	
XX	XX		
PA	(CNRS) CENT NAT RECH SCI.		
XX	XX		
PI	Gronenborn B;		
XX	XX		
DR	WPI; 1996-179947/18.		
XX	XX		
XX	N-PSDB; T12904.		
PT	Prodn. of virus-resistant transgenic plants - using mutated genomic		
XX	sequence from phytopathogenic DNA virus		
PS	Disclosure: Fig 13; 93pp; French.		
XX	XX		
CC	Mutation of consensus amino acids in the NTP-binding site of		
CC	geminivirus Rep protein is used to produce replication deficient		
CC	viruses. The mutated viral nucleic acid is used for producing		
CC	transgenic plants that are resistant to, or tolerant of, the native		
CC	virus. The present sequence is a mutant form of the Rep (Cp) protein		
CC	originally present in the genome of the geminivirus (Cp) virus		
CC	(STVLCV) in which the wild-type Lys227 residue has been changed to an		
CC	Ala residue; transgenic Nicotiana benthamiana plants generated by an		
CC	transformation with the mutated virus were found to be resistant to		
CC	STVLCV, i.e. the mutation results in a dominant negative phenotype.		
XX	XX		
SQ	Sequence	359 AA;	
	Query Match	59.4%; Score 215; DB 17; Length 359;	
	Best Local Similarity	55.1%; Pred. No. 1.4e-18;	
	Matches	38; Conservative 12; Mismatches 19; Indels 0; Gaps	
Qy	2	LWMEFQVDGAAGGCGTSSNDAAEALNASSKEEALQITREKIPEKYLQFPHNSNLD 61	
Db	111	lwgvtqqlgrsrggggtandysakainagksqalvikeleaprdyvhfhmnsnld 170	
		.....	.....
Qy	62	DRIFDKTP	70
		.....	.....
Db	171	kvfqppap	179
		.....	.....

```

R88872
ID R88872 standard; Protein: 359 AA.
XX
XX R88872;
XX
DE 07-NOV-1996 (first entry)
XX
XX Sardinian tomato yellow leaf curl virus mutated Rep protein (K227R).
XX
XX Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
XX modification; mutation; viral replication; deficient; inhibition;
XX etc. Sardinian tomato yellow leaf curl virus;
XX Sardinian isolate, SYLGV; transgenic plant, P-loop, CI protein;
XX A11 protein.
XX
XX Sardinian tomato yellow leaf curl virus.
XX
XX Key Location/Qualifiers
FH Misc-difference 227
FT /note= "wild-type Lys has been replaced by Arg"
XX
XX W09608573-A1.
XX
XX 21-MAR-1996.
XX
XX 15-SEP-1995; 95WO-FR01192.
XX
XX 15-SEP-1994; 94FR-0011040.
XX
XX (CNRS ) CENT NAT RECH SCI.
XX
XX Gronenborn B;
XX
XX WPI; 1996-179947/18.
XX
XX N-PSDB; T12906.
XX
XX Prodn. of virus-resistant transgenic plants - using mutated genomic
XX sequence from phytopathogenic DNA virus
XX
XX Disclosure: Fig 13; 93pp; French.
XX
XX Mutation of consensus amino acids in the NTP-binding site of
XX geminivirus protein is used to produce replication deficient
XX transgenic plants that are resistant to or tolerant of the
XX virus. The present sequence is a mutant form of the Rep (or CI)
XX protein from the Sardinian isolate of tomato yellow leaf curl virus
XX (SYLGV) in which the wild-type Lys227 residue has been changed to
XX an Arg residue; transgenic Nicotiana benthamiana plants generated by
XX transformation with the mutated virus were not resistant to SYLGV.
XX The plants transformed with a virus in which Lys227 had been
XX replaced by Ala were found to be resistant.
XX
XX Sequence 359 AA:

Query Match 59.4%; Score 215; DB 17; Length 359;
Best Local Similarity 55.1%; Pred. No. 1.4e-18;
Matches 39; Conservative 12; Mismatches 19; Indels 0; Gaps 0;

QY 2 LWGCEPVDGAAAGCGTSDNAAALNASSKEPALQIIREKIPEKYLQFHNLSNL 61
DB 111 LWGTFQIGRSGRGQGTQDAYKAINAGSKQALDIKELAPDXYVHFMNSL 170
QY 62 RIFDKTPPEP 70
DB 171 KVFQVPPPEP 179

RESULT 9
ID P70407 standard; Protein: 353 AA.
XX
XX

```

```

P70407;
XX
XX 02-MAY-1991 (first entry)
XX
XX ORF 4 gene product of Bean Golden Mosaic virus.
XX
XX Plant vector.
XX
XX Bean golden mosaic virus.
XX
XX JP61257186-A.
XX
XX 14-NOV-1986.
XX
XX 10-MAY-1965; 85JP-0098108.
XX
XX 10-MAY-1985; 85JP-0098108.
XX
XX (TEIJ ) TEIJIN KK.
XX
XX WPI; 1987-150662/23.
XX
XX N-PSDB; N70630.
XX
XX New DNA and hybrid DNA - used for recombinant vector of plants.
XX
XX Disclosure: Fig 6; 24pp; Japanese.
XX
XX The sequence encoding this protein may be taken along with the -a
XX DNA sequence and a suitable resistance gene, and used to create a
XX recombinant vector.
XX
XX See also N70623.
XX
XX Sequence 353 AA:

Query Match 58.8%; Score 213; DB 8; Length 353;
Best Local Similarity 57.1%; Pred. No. 2.4e-18;
Matches 40; Conservative 14; Mismatches 16; Indels 0; Gaps 0;

QY 1 LTVKCEPVDGAAAGCGTSDNAAALNASSKEPALQIIREKIPEKYLQFHNLSNL 60
DB 110 LIEWGQFQVGRSGRGQGTQDAYKAINADSIESALTILKEEQPKDYVIGHNIRSL 169
QY 61 RIFDKTPPEP 70
DB 170 ERIFKTPPEP 179

RESULT 10
W34336
ID W34336 standard; Protein: 361 AA.
XX
XX W34336;
XX
XX 27-APR-1998 (first entry)
XX
XX Tomato mottle virus AC1 protein.
XX
XX Geminivirus; TOMOV; AC1 gene; transdominant mutation;
XX transgenic plant; disease resistance.
XX
XX Tomato mottle virus isolate Florida.
XX
XX W09739110-A1.
XX
XX 23-OCT-1997.
XX
XX 15-APR-1997; 97WO-US06300.
XX
XX 16-APR-1996; 96US-0015517.
XX
XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.
XX
XX (WISC ) WISCONSIN ALUMNI RES FOUND.

```

XX Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;  
 PI WPI: 1997-526447/48.  
 DR N-PSDB: T93284.  
 XX  
 XX Transgenic plants expressing geminivirus AC1 and CI wild-type and  
 PT mutant genes - have increased resistance to geminivirus infection  
 PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean  
 TT golden mosaic geminivirus  
 XX  
 XX Example 3.4: Page 57-59; 132pp; English.  
 XX  
 XX This protein comprises the wild-type AC1 protein of tomato mottle  
 CC virus (ToMoV), a geminivirus that has a bipartite genome. The AC1  
 CC gene (see T93294), a geminivirus that has a bipartite genome. The AC1  
 CC the two genomic components, DNA-A and DNA-B. The AC1 protein has a  
 CC DNA binding site specific to the DNA-A common region, a DNA nicking  
 CC producton and an NTP-binding activity. The invention involves  
 CC wild-type or mutant sequences that negatively interfere with traps  
 CC with geminiviral replication during infection. Such transgenic  
 CC plants are resistant to viral infection. The AC1/CI genes are  
 CC especially from ToMoV, tomato yellow leaf curl virus or bean golden  
 CC mosaic geminivirus (see T93282-93) and encode polypeptides (see  
 CC W34324-35) that have mutations in the highly conserved DNA-nicking  
 XX domain and/or the NTP-binding domain.  
 XX  
 XX Sequence 361 AA;  
 SQ

Query Match 56.8%; Score 213; DB 18; Length 361;  
 Best Local Similarity 54.3%; Pred. No. 2.5e-18;  
 Matches 38; Conservative 17; Mismatches 15; Indels 0; Gaps 0;  
 OY 1 TLVKEFVDGAAAGCQTSNDAAAEALNASKEEALQITREKIPKYLFOFHINLSNL 60  
 DB 110 tlewdgfqdgisrarg9ggsandysakalnasvgsalavireeqkdfvlqnhirsnl 169  
 OY 61 DRIFDKTPPEP 70  
 DB 170 erifakapep 179

RESULT 11  
 ID W34324 standard; Protein: 361 AA.  
 XX W34324;  
 XX W34324;  
 XX 27-APR-1998 (first entry)  
 XX Tomato mottle virus AC1 mutant ToMoV-AC1dml.  
 DE Geminivirus; ToMoV-AC1dml; AC1 gene; transdominant mutation;  
 XX transgenic plant; disease resistance.  
 XX Tomato mottle virus isolate Florida.  
 OS Synthetic.  
 XX W09739110-A1.  
 XX W09739110-A1.  
 XX 23-OCT-1997.  
 XX 15-APR-1997; 97WO-US06300.  
 XX 16-APR-1996; 96US-0015517.  
 XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.  
 XX (WISC ) WISCONSIN ALUMNI RES FOUND.  
 XX Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;  
 PI WPI: 1997-526447/48.  
 DR N-PSDB: T93283.

DR WPI: 1997-526447/48.  
 XX N-PSDB: T93282.  
 XX Transgenic plants expressing geminivirus AC1 and CI wild-type and  
 PT mutant genes - have increased resistance to geminivirus infection  
 PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean  
 TT golden mosaic geminivirus  
 XX  
 XX Example 3.4: Page 60-62; 132pp; English.  
 XX  
 XX This protein comprises a transdominant lethal mutant, designated  
 CC ToMoV-AC1dml, of tomato mottle virus (ToMoV) AC1 protein (see  
 CC T93294) and a transdominant lethal mutant, designated ToMoV-  
 CC tomato mottle virus (ToMoV) and conserved mutations in its  
 CC NTP-binding domains. The AC1 gene (see also T93294) must be  
 CC expressed for efficient replication of the two genomic components,  
 CC DNA-A and DNA-B, of the bipartite ToMoV genome. The invention  
 CC involves production of transgenic plants containing DNA comprising  
 CC geminivirus AC1 or CI wild-type or mutant sequences that negatively  
 CC interfere in trans with geminiviral replication during infection.  
 CC The AC1/CI genes are especially from ToMoV, tomato yellow leaf  
 CC curl virus or bean golden mosaic geminivirus (see T93282-93) and encode  
 CC polypeptides (see W34324-35) that have mutations in the highly  
 CC conserved DNA-nicking and/or NTP-binding domains.  
 XX  
 XX Sequence 361 AA;  
 SQ

Query Match 58.8%; Score 213; DB 18; Length 361;  
 Best Local Similarity 54.3%; Pred. No. 2.5e-18;  
 Matches 38; Conservative 17; Mismatches 15; Indels 0; Gaps 0;  
 OY 1 TLVKEFVDGAAAGCQTSNDAAAEALNASKEEALQITREKIPKYLFOFHINLSNL 60  
 DB 110 tlewdgfqdgisrarg9ggsandysakalnasvgsalavireeqkdfvlqnhirsnl 169  
 OY 61 DRIFDKTPPEP 70  
 DB 170 erifakapep 179

RESULT 12  
 ID W34325 standard; Protein: 361 AA.  
 XX W34325;  
 XX W34325;  
 XX 27-APR-1998 (first entry)  
 XX Tomato mottle virus AC1 mutant ToMoV-AC1dml.  
 DE Geminivirus; ToMoV-AC1dml; AC1 gene; transdominant mutation;  
 XX transgenic plant; disease resistance.  
 XX Tomato mottle virus isolate Florida.  
 OS Synthetic.  
 XX W09739110-A1.  
 XX W09739110-A1.  
 XX 23-OCT-1997.  
 XX 15-APR-1997; 97WO-US06300.  
 XX 16-APR-1996; 96US-0015517.  
 XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.  
 XX (WISC ) WISCONSIN ALUMNI RES FOUND.  
 XX Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;  
 PI WPI: 1997-526447/48.  
 DR N-PSDB: T93283.



XX Transgenic plants expressing geminivirus AC1 and CI wild-type and  
 PT mutant genes - have increased resistance to geminivirus infection  
 PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean  
 PT golden mosaic geminivirus  
 XX  
 XX Example 3.5; Page 64-65; 132pp; English.  
 XX  
 CC This protein comprises a transdominant lethal mutant, designated  
 CC ToMoV-AC1dlm23, of tomato mottle virus (ToMoV) AC1 protein (see  
 CC W34326). It is encoded by a mutant AC1 gene (see T93283) of  
 CC ToMoV-AC1dlm23. The AC1 gene carries 2 mutations in an  
 CC NTP-binding domain. The AC1 gene (see also T93284) must be  
 CC expressed for efficient replication of the two genomic components,  
 CC DNA-A and DNA-B, of the bipartite ToMoV genome. The mutant invention  
 CC involves production of transgenic plants containing DNA comprising  
 CC geminivirus AC1 or CI wild-type or mutant sequences that negatively  
 CC interfere in trans with geminiviral replication during infection. The  
 CC such transgenic plants are resistant to viral infection. The  
 CC AC1/CI genes are especially from ToMoV, tomato yellow leaf curl  
 CC virus or bean golden mosaic geminivirus (see T93282-93) and encode  
 CC polypeptides (see W34324-35) that have mutations in the highly  
 CC conserved DNA-nicking and/or NTP-binding domains.  
 XX  
 SQ Sequence 361 AA;

Query Match 58.8%; Score 213; DB 18; Length 361;  
 Best Local Similarity 54.3%; Pred. No. 2,5e-18;  
 Matches 38; Conservative 17; Mismatches 15; Indels 0; Gaps 0;  
 QY 1 TLVKGFEQVGDGAAGGCGTSDNAAAFALNASSKEALQLIREKIPKYLGFPHNLNSHL 60  
 Db 110 TLVKGFEQVGDGAAGGCGTSDNAAAFALNASSKEALQLIREKIPKYLGFPHNLNSHL 60  
 QY 61 DRIPDKTPEP 70  
 Db 170 erifakepep 179

RESULT 13  
 ID W34326 standard; Protein; 361 AA.  
 XX  
 AC W34326;  
 XX  
 DT 27-APR-1998 (first entry)  
 XX  
 DE Tomato mottle virus AC1 mutant ToMoV-AC1dlm23.  
 KW Geminivirus; ToMoV-AC1dlm23; AC1 gene; transdominant mutation;  
 XX transgenic plant; disease resistance.  
 OS  
 OS Tomato mottle virus isolate Florida.  
 XX  
 XX W09739110-A1.  
 XX  
 PD 23-OCT-1997.  
 XX  
 PF 15-APR-1997; 97WO-0506300.  
 XX  
 XX 16-APR-1996; 96US-0015517.  
 PR (SEM-) SEMINIS VEGETABLE SEEDS INC.  
 PA (WISC ) WISCONSIN ALUMNI RES FOUND.  
 XX  
 PI Ahlquist PG, Hanson SF, Lau HT, Maxwell DP, Stout JT;  
 XX WPI: 1997-526447/48.  
 DR N-PSDB: T93284-END.  
 XX  
 XX Transgenic plants expressing geminivirus AC1 and CI wild-type and

PT mutant genes - have increased resistance to geminivirus infection  
 PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean  
 PT golden mosaic geminivirus  
 XX  
 XX Example 3.6; Page 67-69; 132pp; English.  
 XX  
 CC This protein comprises a transdominant lethal mutant, designated  
 CC ToMoV-AC1dlm23, of tomato mottle virus (ToMoV) AC1 protein (see  
 CC W34336). It is encoded by a mutant AC1 gene (see T93284) of  
 CC ToMoV-AC1dlm23. The AC1 gene carries 2 mutations in an  
 CC NTP-binding domain. The AC1 gene (see also T93284) must be  
 CC expressed for efficient replication of the two genomic components,  
 CC DNA-A and DNA-B, of the bipartite ToMoV genome. The mutant invention  
 CC involves production of transgenic plants containing DNA comprising  
 CC geminivirus AC1 or CI wild-type or mutant sequences that negatively  
 CC interfere in trans with geminiviral replication during infection. The  
 CC such transgenic plants are resistant to viral infection. The  
 CC AC1/CI genes are especially from ToMoV, tomato yellow leaf curl  
 CC virus or bean golden mosaic geminivirus (see T93282-93) and encode  
 CC polypeptides (see W34324-35) that have mutations in the highly  
 CC conserved DNA-nicking and/or NTP-binding domains.  
 XX  
 SQ Sequence 361 AA;

Query Match 58.8%; Score 213; DB 18; Length 361;  
 Best Local Similarity 54.3%; Pred. No. 2,5e-18;  
 Matches 38; Conservative 17; Mismatches 15; Indels 0; Gaps 0;  
 QY 1 TLVKGFEQVGDGAAGGCGTSDNAAAFALNASSKEALQLIREKIPKYLGFPHNLNSHL 60  
 Db 110 TLVKGFEQVGDGAAGGCGTSDNAAAFALNASSKEALQLIREKIPKYLGFPHNLNSHL 60  
 QY 61 DRIPDKTPEP 70  
 Db 170 erifakepep 179

RESULT 14  
 ID P70562 standard; Protein; 361 AA.  
 XX  
 AC P70562;  
 XX  
 DT 30-APR-1991 (first entry)  
 XX  
 DE Product of ORF 4 from MYMV complementary strand (b).  
 XX  
 KW Geminivirus.  
 XX Mungbean yellow mosaic virus.  
 XX  
 PN JP62126982-A.  
 XX  
 PD 09-JUN-1987.  
 XX  
 PF 28-NOV-1985; 85JP-0266080.  
 XX  
 XX 28-NOV-1985; 85JP-0266080.  
 XX  
 XX (TEIJ ) TEIJIN KK.  
 PA  
 XX WPI: 1987-196308/28.  
 DR N-PSDB: N70895.  
 XX  
 XX Novel DNA and hybrid DNA useful - as vector for recombinant work  
 PT of plant gene.  
 XX  
 XX Disclosure: Fig 8; 21pp; Japanese.  
 XX  
 CC The sequence is encoded by ORF 4 which occurs on the complementary  
 CC strand of the (b) molecule of the geminivirus.  
 CC See also P70559-P70567.

```

XX      Sequence      361 AA;
SQ

Query Match          57.0%; Score 206.5; DB 8; Length 361;
Best Local Similarity 63.1%; Pred. No. 1.6e-17;
Matches 41; Conservative 10; Mismatches 13; Indels 1; Gaps 1;

QY      5 GFEQVDCGAAAGGCQTSDNDAAEALNASKEEALQIIREKIPKYLQFQHNLSNLDRI 64
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      114 gsfqvgdgrsar9gkgsandayaealnsgsklqalnllrekapkeyllqfhnlnsnl 173
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      65 -DKTP 68
      | | | | |
DB      174 sdevp 178

RESULT 15
W56495 standard; Protein: 362 AA.
XX      W56495:
AC      W56495:
XX      11-AUG-1998 (first entry)
XX      Tobacco leaf curling virus gene protein Cl.
DE      DE
XX      Tobacco leaf curling virus gene; TICV; promoter; Cl protein.
XX      Tobacco leaf curling virus.
OS      JP10070982-A.
XX      17-MAR-1998.
PD      PD
XX      30-AUG-1996; 96JP-0230394.
XX      30-AUG-1996; 96JP-0230394.
PR      (NORQ ) NORINSUISANSHO KYUSHU NOGYO SHIKENJO.
XX      WPI; 1998-233630/21.
XX      Tobacco leaf curling virus gene - useful for inserting into vectors
PT      for expression in, e.g. tomato plants
XX      Claim 5; Fig 5; 9pp; Japanese.
XX      This sequence represents the Cl protein encoded by the tobacco leaf
CC      curling virus (TICV) gene of the invention. TICV gene or its promoter can
CC      be inserted into a vector for expression in plants, e.g. tobacco and
CC      tomato. This sequence is believed to be encoded by the TICV gene shown in
CC      V29761.
XX      Sequence      362 AA;
SQ

Query Match          55.5%; Score 201; DB 19; Length 362;
Best Local Similarity 60.6%; Pred. No. 7.4e-17;
Matches 42; Conservative 11; Mismatches 15; Indels 16; Gaps 1;

QY      2 LVNGEFDVCGAAGGCQTSDNDAAEALNASKEEALQIIREKIPKYLQFQHNLSNLD 61
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      111 ldfgvgdgrsarccqandayaealnsgskalnllrekapkfvlqfhnlnsnld 170
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      62 RT-----FKTTP 69
      | | | | |
DB      171 rifaplevfcvffsssfddvpe 194

```

Search completed: February 3, 2001, 02:15:26  
Job time: 3433 sec

Run on: February 3, 2001, 02:22:50 ; Search time 180.33 Seconds  
(without alignments)  
45.497 Million cell updates

TLVNGEFOVDGRSARGGCOT.....FQFHNLSNLDRIFDKTPPEP 70

iaipop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 374700

000000000000 1016 CLOSING PAGE \*\*\* 000000000000

FOR FURTHER INFORMATION CONTACT:

```
3: sp_fungi:*
```

7: sp\_mhc:★

```
11: sp_rodent:*
```

[illegible]

.....

No.	Score	Match Length	DB	ID	Des
-----	-------	--------------	----	----	-----

3	264	73.7	226	12	Q9WHF6
3	264	73.7	226	12	Q9WHF6

7	260	72.6	233	12	Q9YLA4	Q9YLA4

11	250	69.8	190	12	Q9W827	Q9
----	-----	------	-----	----	--------	----

Year	15	249	69.6	208	12	Q9Z0A0	Q9
2000	15	249	69.6	208	12	Q9Z0A0	Q9

19	245	68.4	203	12	Q9Z083	Q9
----	-----	------	-----	----	--------	----

## ALIGNMENTS

Query Match	77.1%	Score 276;	DB 12;	Length 226;
Best Local Similarity	77.1%;	Pred. No. 3.1e-24;		
Matches 54;	Conservative	4;	Mismatches 12;	Indels 0; Gaps 0;

2y 61 DRIFDKTPEP 70  
|||||  
171 DRIFAKAPEP 180

RESULT 2  
99QS55

```

ID 09Q555; PRELIMINARY; PRT: 364 AA.
AC 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE REPLICATION INITIATION PROTEIN AC1.
GN AC1.
OS Sweet potato leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=100755;
RP SEQUENCE FROM N.A.
RA Detraction of a geminivirus infecting sweet potato in the United States.*;
RL Plant Dis. 82:1253-1257(1998).
DR ENBL; AF104036; AA047173.1; -.
DR INTERPRO; IPR001191; -.
DR INTERPRO; IPR001301; -.
DR PRINTS; PR00227; GEMCOATALL.
DR PRINTS; PR00228; GEMCOATCLVL.
SQ SEQUENCE 364 AA; 40680 MW; 5F79752431A09D6E CRC64;

Query Match
Best Local Similarity 75.7%; Score 271; DB 12; Length 364;
Matches 54; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSARGCGQTSNDAAALNASSAAALQIIREKIPKYLFOFHNLNSNL 60
DB 110 TLVWGEFQVDRSARGCGQTSNDAAALNASSAAALQIIREKIPKYLFOFHNLVSNL 169

QY 61 DRIFOKTP 68
DB 170 DRIFSTPP 177

RESULT 3
ID 09MHF6 PRELIMINARY; PRT: 226 AA.
AC 01-MAY-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DE 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN REP.
OS tomato mild mottle geminivirus.
OC Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.
OX NCBI_TaxID=92343;
RP SEQUENCE FROM N.A.
RA Khalila M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M., Maxwell D.P.;
RT "Molecular characterization and DNA-based detection methods for vegetable-infecting geminiviruses in Central America";
DR ENBL; AF131073; AA033471.1; -.
DR INTERPRO; IPR001301; -.
DR INTERPRO; IPR001301; -.
DR PFAM; PF00799; Gemin_Ali; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR PRINTS; PR00228; GEMCOATCLVL.
DR ANINTER 226
SQ SEQUENCE 226 AA; 25941 MW; 2EA4116712871A23 CRC64;

Query Match
Best Local Similarity 73.7%; Score 264; DB 12; Length 226;
Matches 49; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSARGCGQTSNDAAALNASSAAALQIIREKIPKYLFOFHNLNSNL 60

```

```

DB 111 TLVWGEFQVDRSARGCGQTSNDAAALNASSAAALQIIREKIPKYLFOFHNLSSNL 170
QY 61 DRIFOKTPPP 70
DB 171 DRIFAKAPEP 180

RESULT 4
ID 067574 PRELIMINARY; PRT: 361 AA.
AC 01-MAY-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DE 01-JUN-2000 (TRENBLrel. 14, Last annotation update)
DE PUTATIVE REPLICATIVE PROTEIN.
GN Ali.
OS Bean golden mosaic virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10839;
RP SEQUENCE FROM N.A.
RA Gilbertson R.L., Farla J.C., Hanson S.F., Morales F.J., Ahlquist P.G., Maxwell D.P., Russell D.R.;
RL Phytopathology 81:980-985(1991).
RN [2]
RN SEQUENCE FROM N.A.
RA Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Farla J.C., Plant Dis. 75:336-342(1991).
RN [3]
RN SEQUENCE FROM N.A.
RA Gilbertson R.L., Ahlquist P.G., Maxwell D.P.;
RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.
DR ENBL; M8686; AA046312.1; -.
DR INTERPRO; IPR001301; -.
DR INTERPRO; IPR001301; -.
DR PFAM; PF00799; Gemin_Ali; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR PRINTS; PR00228; GEMCOATCLVL.
DR PRODOM; PD000736; -.
SQ SEQUENCE 361 AA; 41041 MW; 0094C7ACAF06B788 CRC64;

Query Match
Best Local Similarity 73.3%; Score 262; DB 12; Length 361;
Matches 49; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

QY 4 WGEFQVDRSARGCGQTSNDAAALNASSAAALQIIREKIPKYLFOFHNLMLORI 63
DB 113 WGEFQVDRSARGCGQTSNDAAALNASSAAALQIIREKIPKYLFOFHNLMLORI 172

QY 64 FDKTPEP 70
DB 173 FTKAPDP 179

RESULT 5
ID 068975 PRELIMINARY; PRT: 149 AA.
AC 01-MAY-1997 (TRENBLrel. 03, Created)
DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)
DE 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN AC1.
OS Coptidium golden mosaic geminivirus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=51676;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN-JAMAICAN;
RA ROYE M.E.;

```

RL Thesis (1996), Biochemistry, University of the West Indies, Jamaica.  
 DK EMBL: U75278; AAB36519.1; -  
 DR INTERPRO: IPR001191; -  
 DR NCBI\_TaxID: 501; 1; All: 1.  
 DR PFAM: PF00759; GEMCOATL1.  
 DR PRINTS: PRO0227; GEMCOATL1.  
 DR PRINTS: PRO0228; GEMCOATL1.  
 FT NON\_TER 1 149  
 FT NON\_TER 1 149  
 SQ SEQUENCE 149 AA; 16785 MW; E4CF5EDC49CD508 CRC64;

Query Match 72.6%; Score 260; DB 12; Length 149; -  
 Best Local Similarity 68.6%; Pred. No. 1.4e-22;  
 Matches 48; Conservative 12; Mismatches 10; Indels 0; Gaps 0;

Qy 1 TLWGFVGVGARGCGQTSNDAAAALAAASSAAALQIIREKIPKYLIFQPHNLNSL 60  
 Db 52 TLWGFVGVGARGCGQTSNDAAAALAAASSAAALQIIREKIPKYLIFQPHNLNSL 111  
 Qy 61 DRIFDKTPEP 70  
 Db 112 DRIFMDPEP 121

RESULT 6  
 Q90BL SQ0B1 PRELIMINARY; PRT; 225 AA.  
 AC Q90BL;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)  
 DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).  
 GN REP.  
 OS Cowpea golden mosaic geminivirus.  
 OS Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 OC NCBI\_TaxID: 63263;  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-CGMV-BR;  
 RA Faria J.C.;  
 RT Partial nucleotide sequence of cowpea golden mosaic geminivirus from submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 RL EMBL: AF186708; AAF06318.1; -  
 DR INTERPRO: IPR001191; -  
 DR INTERPRO: IPR001301; -  
 DR PFAM: PF00799; Gemini\_A11; 1.  
 DR PRINTS: PRO0227; GEMCOATL1.  
 DR PRINTS: PRO0228; GEMCOATL1.  
 FT NON\_TER 1 25766 MW; 1089CH6BD8D1565D CRC64;  
 SQ SEQUENCE 225 AA; 25766 MW; 1089CH6BD8D1565D CRC64;

Query Match 72.6%; Score 260; DB 12; Length 225;  
 Best Local Similarity 73.1%; Pred. No. 2.2e-22;  
 Matches 49; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

Qy 4 WGFQVGVGARGCGQTSNDAAAALAAASSAAALQIIREKIPKYLIFQPHNLNSLDR1 63  
 Db 113 WGFQVGVGARGCGQTSNDAAAALAAASSAAALQIIREKIPKYLIFQPHNLNSLDR1 172  
 Qy 64 FDKTPEP 70  
 Db 173 FRKPPPEP 179

RESULT 7  
 Q9VLA4 PRELIMINARY; PRT; 233 AA.  
 ID Q9VLA4  
 AC Q9VLA4;  
 DT 01-MAY-1999 (TReMBLrel. 10, Created)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)

DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)  
 DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).  
 GN REP.  
 OS Macrotellium golden mosaic geminivirus.  
 OS Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 OC NCBI\_TaxID: 51676;  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-JAMAICA STRAIN 1;  
 RA Roye M.E.;  
 RT Genetic diversity and phylogeny of whitefly-transmitted geminiviruses from Jamaica.  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-JAMAICA STRAIN 1;  
 RA Roye M.E., McLaughlin W.A., Maxwell D.P.;  
 RT Molecular characterization of two distinct geminiviruses infecting M. atyloides from Jamaica.  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 RL EMBL: AF098940; AAD17850.1; -  
 DR INTERPRO: IPR001191; -  
 DR INTERPRO: IPR001301; -  
 DR PFAM: PF00799; Gemini\_A11; 1.  
 DR PRINTS: PRO0227; GEMCOATL1.  
 DR PRINTS: PRO0228; GEMCOATL1.  
 FT NON\_TER 233 25355 MW; AA490AFD2166A02 CRC64;  
 SQ SEQUENCE 233 AA; 26355 MW; AA490AFD2166A02 CRC64;

Query Match 72.6%; Score 260; DB 12; Length 233;  
 Best Local Similarity 68.6%; Pred. No. 2.2e-22;  
 Matches 48; Conservative 12; Mismatches 10; Indels 0; Gaps 0;

Qy 1 TLWGFVGVGARGCGQTSNDAAAALAAASSAAALQIIREKIPKYLIFQPHNLNSL 60  
 Db 110 TLWGFVGVGARGCGQTSNDAAAALAAASSAAALQIIREKIPKYLIFQPHNLNSL 169  
 Qy 61 DRIFDKTPEP 70  
 Db 170 DRIFMDPEP 179

RESULT 8  
 Q98693 PRELIMINARY; PRT; 185 AA.  
 ID Q98693  
 AC Q98693;  
 DT 01-FEB-1997 (TReMBLrel. 02, Created)  
 DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
 DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).  
 GN AC1  
 OS sida golden mosaic virus.  
 OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 OC NCBI\_TaxID: 51034;  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-JAMAICA;  
 RA McLaughlin W.A., Nakhla N.K., Maxwell D.P.;  
 RL Plant Dis., 81:1251-1258(1997).  
 DR EMBL: U67926; AAB97865.1; -  
 DR INTERPRO: IPR001191; -  
 DR INTERPRO: IPR001301; -  
 DR PFAM: PF00799; Gemini\_A11; 1.  
 DR PRINTS: PRO0227; GEMCOATL1.  
 DR PRINTS: PRO0228; GEMCOATL1.  
 FT NON\_TER 185 185  
 SQ SEQUENCE 185 AA; 20975 MW; 3913850A025A5EE1 CRC64;

Query Match 72.3%; Score 259; DB 12; Length 185;  
 Best Local Similarity 68.6%; Pred. No. 2.3e-22;

Matches 48; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

OY 1 TLWGEFQVDRSGRGCCOTSNDAAALAAASSAAALQIIREKIPKYLFOFHNSNL 60  
 DB 89 TLWGTFTQDGRSARGGQANDANAKAALSGTKEDAKIIREKIPKYLFOFHNSNL 148  
 QY 61 DRIFKPTPEP 70  
 DB 149 DRIFSKPPEP 158

## RESULT 9

Q92089 ID Q92089 PRELIMINARY; PRT: 190 AA.  
 AC Q92089;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE C1 AND C4 GENES, CLONE YOKOHAMA3-1.  
 DE PARTIAL AND COMPLETE CDS (FRAGMENT).  
 GN C1  
 OS Tobacco leaf curl virus.  
 OS Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=67762;  
 RN [1]  
 RP STRAIN-YOKOHAMA3;  
 RC Ooi K., Ohshita S., Ishii I., Yahara T.;  
 RT "Molecular phylogeny of geminiviruses infecting wild plants in Japan."  
 RL J. Plant Res. 110:247-257(1997).  
 DR EMBL: AB001315; BAA34033.1; -;  
 DR INTERPRO: IPR001191; -;  
 DR FRAM: PF00799; GeminLALI; 1.  
 DR PRINTS: PR00227; GEMCOATLVL.  
 DR NON\_TER 1  
 FT NON\_TER 1  
 SEQUENCE 190 AA; 21432 MW; AAC093D1D1610FAD CRC64;

## Query Match

Best Local Similarity 70.1%; Score 251; DB 12; Length 190;  
 Matches 50; Conservative 8; Mismatches 11; Indels 16; Gaps 1;

OY 1 TLWGEFQVDRSGRGCCOTSNDAAALAAASSAAALQIIREKIPKYLFOFHNSNL 60  
 DB 85 TLWGTFTQDGRSARGGQANDANAKAALSGTKEDAKIIREKIPKYLFOFHNSNL 144  
 QY 61 DRI-----FDKTPPE 69  
 DB 145 DRIFAPPLEVFCVFSSFDQVPE 169

## RESULT 10

Q92084 ID Q92084 PRELIMINARY; PRT: 190 AA.  
 AC Q92084;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE C1 AND C4 GENES, CLONE YOKOHAMA5-2.  
 DE PARTIAL AND COMPLETE CDS (FRAGMENT).  
 GN C1  
 OS Tobacco leaf curl virus.  
 OS Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=67762;  
 RN [1]  
 RP STRAIN-YOKOHAMA5;  
 RC Ooi K., Ohshita S., Ishii I., Yahara T.;  
 RT "Molecular phylogeny of geminiviruses infecting wild plants in Japan."  
 RL J. Plant Res. 110:247-257(1997).

DR EMBL: AB001315; BAA34033.1; -;  
 DR INTERPRO: IPR001191; -;  
 DR FRAM: PF00799; GeminLALI; 1.  
 DR PRINTS: PR00227; GEMCOATLVL.  
 DR NON\_TER 1  
 FT NON\_TER 1  
 SEQUENCE 190 AA; 21444 MW; AAC1C2943E3F01AD CRC64;

## Query Match

Best Local Similarity 70.1%; Score 251; DB 12; Length 190;  
 Matches 50; Conservative 8; Mismatches 11; Indels 16; Gaps 1;

OY 1 TLWGEFQVDRSGRGCCOTSNDAAALAAASSAAALQIIREKIPKYLFOFHNSNL 60  
 DB 85 TLWGTFTQDGRSARGGQANDANAKAALSGTKEDAKIIREKIPKYLFOFHNSNL 144  
 QY 61 DRI-----FDKTPPE 69  
 DB 145 DRIFAPPLEVFCVFSSFDQVPE 169

## RESULT 11

Q9H827 ID Q9H827 PRELIMINARY; PRT: 190 AA.  
 AC Q9H827; 1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE C1 PROTEIN (FRAGMENT).  
 GN C1  
 OS Tobacco leaf curl virus.  
 OS Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=67762;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-GORAI;  
 RA Ooi K., Ohshita S., Ishii I., Yahara T.;  
 RT "Molecular phylogeny of geminiviruses infecting wild plants in Japan."  
 RL J. Plant Res. 110:247-257(1997).  
 DR EMBL: AB001303; BAA34010.1; -;  
 DR INTERPRO: IPR001191; -;  
 DR FRAM: PF00799; GeminLALI; 1.  
 DR PRINTS: PR00227; GEMCOATLVL.  
 DR NON\_TER 1  
 FT NON\_TER 1  
 SEQUENCE 190 AA; 21444 MW; 93C3742A9BD87EB CRC64;

## Query Match

Best Local Similarity 69.8%; Score 250; DB 12; Length 190;  
 Matches 50; Conservative 8; Mismatches 11; Indels 16; Gaps 1;

OY 1 TLWGEFQVDRSGRGCCOTSNDAAALAAASSAAALQIIREKIPKYLFOFHNSNL 60  
 DB 85 TLWGTFTQDGRSARGGQANDANAKAALSGTKEDAKIIREKIPKYLFOFHNSNL 144  
 QY 61 DRI-----FDKTPPE 69  
 DB 145 DRIFAPPLEVFCVFSSFDQVPE 169

## RESULT 12

Q920C4 ID Q920C4 PRELIMINARY; PRT: 208 AA.  
 AC Q920C4;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)

```

DE C1 AND C4 GENES; CLONE ABUR3-1, PARTIAL
GW C1.
OS tobacco leaf curl virus
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=67762;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ABUR3;
RA Col K., Ohshita S., Ishii I., Yahara T.;
RT "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL J. Plant Res. 110:247-257(1997).
DR EMBL; AB001294; BAA33992.1; -.
DE C1 AND C4 GENES; CLONE AMG-1(815A)
DR INTERPRO; IPR001191; -.
DE PARTIAL AND COMPLETE CDS (FRAGMENT).
DR PFAM; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOTALL1.
DR PRINTS; PR00228; GEMCOTCLVL1.
FT NON_TER 208 208
SQ SEQUENCE 208 AA; 23526 MW; 249C31D8729C72D CRC64;

Query Match 69.8%; Score 250; DB 12; Length 208;
Best Local Similarity 58.8%; Pred. No. 2.8e-21;
Matches 50; Conservative 8; Mismatches 11; Indels 16; Gaps 1;

Qy 1 TLWGEFQVGDGRSANGCCTSDNDAALAAASSAAALQIIREKIPKYLQFQHNLSNU 60
Db 11 TLWGEFQVGDGRSANGCCTSDNDAALAAASSAAALQIIREKIPKYLQFQHNLSNU 60
Qy 61 DRI-----FDKTP 69
Db 162 DRIAPLEVFVCPSSSSFDQVPE 186

RESULT 14
Qy 0920B8 PRELIMINARY; PRT; 208 AA.
ID Q920B8
AC Q920B8
DT 01-MAY-1999 (TEMBLrel. 10, Created)
DR EMBL; AB001296; BAA33996.1; -.
DE C1 AND C4 GENES; CLONE AMG-2(815A)
DE PARTIAL AND COMPLETE CDS (FRAGMENT).
GN C1.
OS tobacco leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=67762;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AMG154;
RA Col K., Ohshita S., Ishii I., Yahara T.;
RT "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL J. Plant Res. 110:247-257(1997).
DR EMBL; AB001297; BAA33998.1; -.
DE C1 AND C4 GENES; CLONE AMG-3(815A)
DR INTERPRO; IPR001191; -.
DE PARTIAL AND COMPLETE CDS (FRAGMENT).
DR PFAM; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOTALL1.
DR PRINTS; PR00228; GEMCOTCLVL1.
FT NON_TER 1 208
SQ SEQUENCE 208 AA; 23472 MW; 629D0DEF7C9956AA CRC64;

Query Match 59.8%; Score 250; DB 12; Length 208;
Best Local Similarity 58.8%; Pred. No. 2.8e-21;
Matches 50; Conservative 8; Mismatches 11; Indels 16; Gaps 1;

Qy 1 TLWGEFQVGDGRSANGCCTSDNDAALAAASSAAALQIIREKIPKYLQFQHNLSNU 60
Db 102 TLWGEFQVGDGRSANGCCTSDNDAALAAASSAAALQIIREKIPKYLQFQHNLSNU 161
Qy 61 DRI-----FDKTP 69
Db 162 DRIAPLEVFVCPSSSSFDQVPE 186

RESULT 15
Qy 0920A0 PRELIMINARY; PRT; 208 AA.
ID Q920A0
AC Q920A0
DT 01-MAY-1999 (TEMBLrel. 10, Created)
DR EMBL; AB001296; BAA33996.1; -.
DE C1 AND C4 GENES; CLONE AMG-1(815A)
DE PARTIAL AND COMPLETE CDS (FRAGMENT).
GN C1.
OS tobacco leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=67762;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MINO2;
RA Col K., Ohshita S., Ishii I., Yahara T.;
RT "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL J. Plant Res. 110:247-257(1997).
DR EMBL; AB001308; BAA34020.1; -.
DE C1 AND C4 GENES; CLONE MINO2-1, PARTIAL
DR INTERPRO; IPR001191; -.
DR INTERPRO; IPR001301; -.
DR PFAM; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOTALL1.

```





GenCore version 4.5

Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:24:24 : Search time 83.07 seconds  
(without alignments)  
27.213 Million cell updates/sec

Title: US-09-289-346a-3

Perfect score: 358

Sequence: 1 TLVWGEFQVGRSGCGCOT.....FQFHNLNDRIFDKTPEP 70

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	332	92.7	352	1 VAL1_TCMV	P03567 tomato gold
2	243	67.9	361	1 VAL1_PYMV	P27258 potato yell
3	240	64.2	380	1 VAL1_CLVN	P14592 cassava lat
4	233	64.2	358	1 VAL1_TYLCV	P14592 cassava lat
5	222	62.0	362	1 VAL1_TYLCV	P36274 tomato yell
6	217	60.6	355	1 VAL1_ABYWV	P21947 abutilon mo
7	216	60.3	349	1 VAL1_PHVUV	Q06923 pepper huas
8	215	60.3	351	1 VAL1_TYLCV	P38609 tomato yell
9	215	60.1	353	1 VAL1_BGVW	P05175 bean golden
10	214	59.8	361	1 VAL1_TCMV	Q06557 tomato mott
11	213	59.7	358	1 VAL1_TCMV	P27258 tomato yell
12	208	58.2	358	1 VAL1_TYLCV	P27258 tomato yell
13	190	53.1	357	1 VAL1_TYLCV	P27258 tomato yell
14	115	32.1	347	1 VAL1_SLCV	P29048 squash leaf
15	69	19.3	577	1 TFPA_DROME	Q05913 drosophila
16	66	18.4	207	1 IL6_MARMO	P35736 marmota mon
17	66	18.4	508	1 CROC_DROME	P23027 drosophila
18	65	18.2	671	1 HROC_DROME	P22810 drosophila
19	63	17.9	671	1 HROC_DROME	Q02308 drosophila
20	60	17.3	1829	1 P11336_DROSOPH	P11336 drosophila
21	60	16.9	883	1 P11336_DROSOPH	P11336 drosophila
22	60	16.8	314	1 PNKB_HUMAN	O95943 homo sapien
23	60	16.8	314	1 PNKB_HUMAN	O35690 mus musculu
24	59.5	16.6	470	1 RHSA_RHIME	Q923r2 rhizobium m
25	59	16.5	379	1 CYB_ASTPE	P23818 asterina pc
26	59	16.5	498	1 CAIR_PIG	P25117 sus scrofa
27	59	16.5	3593	1 HMO_CHICK	P18080 gallus gall
28	59	16.5	1264	1 MOLF_ECOLI	P33345 bacteroides
29	58.5	16.3	1326	1 MOLF_ECOLI	P33345 bacteroides
30	58.5	16.3	1713	1 LNA3_HUMAN	Q16787 homo sapien
31	58	16.2	732	1 CYG4_HUMAN	P13402 homo sapien
32	58	16.2	1081	1 UL52_HSEB	P28962 equine herp
33	58	16.2	1992	1 TR12_HUMAN	Q14669 homo sapien

34 57.5 16.1 303 1 ALB3\_MAIZE  
35 57.5 16.1 557 1 SASB\_ANAPL  
36 57 15.9 387 1 V82\_HPV1  
37 57 15.9 401 1 HME1\_MOUSE  
38 56.5 15.8 664 1 PCH1\_YEAST  
39 56.5 15.8 1205 1 NKCL\_MOUSE  
40 56.5 15.8 1205 1 NKCL\_MOUSE  
41 56 15.6 105 1 RLA2\_LEIBR  
42 56 15.6 398 1 ODO2\_AZOVI  
43 56 15.6 447 1 CDSA\_DROME  
44 56 15.6 582 1 HMO\_DPSTA  
45 56 15.6 1114 1 DEFT\_CHICK

ALIGNMENTS

RESULT 1  
VAL1\_TCMV  
AC P03567 STANDARD: PRT; 352 AA.  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE AL1 PROTEIN.  
GN AC1.  
OS Tomato golden mosaic virus (TCMV).  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
RN  
PE  
PP SEQUENCE FROM N.A.  
RA Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.:  
RT of tomato golden mosaic virus: potential coding regions and regulatory  
RT sequences.  
RL EMBL J. 3:2197-2205(1984).  
CC --- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.ebi-sb.ch/announce/  
CC -- send an email to: ebi-sb-staff@ebi.ac.uk).  
CC EMBL: Q02029; \*: NOT\_ANNOTATED\_CDS.  
DR PIR: A04170; Q0CVL1.  
DR INTERPRO: IPR001191; .  
DR PFAM: PF00759; Gemin1\_AL1; 1.  
DR PRINTS: PR00227; GEMCONTALL1.  
DR PROSITE: PS00226; GEMCONTALLVL.  
KW ATP-binding; 223 230 ATP (POTENTIAL).  
FT NP\_BIND 223 230  
SQ SEQUENCE 352 AA; C33C938B9644B4A4 CRC64:  
92.7% Score 332; DB 1; Length 352;  
Query Match  
Best Local Similarity 92.9%; Pred. No. 6.5e-32;  
Matches 65; Conservative 5; Mismatches 0; Gaps 0;

OY 1 TLVWGEFQVGRSGCGCOTSDAAALAAALAAALQIIRKIPKYLQFQHNLSNL 60  
Db 111 TLVWGEFQVGRSGCGCOTSDAAALAAALAAALQIIRKIPKYLQFQHNLSNL 170  
OY 61 DRIFDKTPEP 70  
Db 171 DRIFDKTPEP 180  
RESULT 2  
VAL1\_PYMV STANDARD: PRT; 361 AA.

P27258;  
 AC DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DE ALL PROTEIN (40.4 KDA protein update)  
 OS Potato yellow mosaic virus (Isolate Venezuela).  
 OS Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 RN PR00227; GEMCOATCLVLI.  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE-91311403; PubMed-1856690;  
 RA Coute R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;  
 RT The nucleotide sequence of the infectious cloned DNA components of  
 RL J. Gen. Virol. 72:1515-1520(1991).  
 CC 1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; D00940; BAA00782.1; -  
 DR J00364; GQCVPT.  
 DR INTERPRO; IPR001301; -  
 DR PFAM; PF00759; GeminLALL; 1.  
 DR PRINTS; PR00227; GEMCOATL1.  
 DR PRINTS; PR00228; GEMCOATCLVLI.  
 DR NP-BINDing. 229  
 FT SEQUENCE 361 AA; 40850 MW; 5627A33BF1264383 CRC64;  
 SQ

Query Match 67.9%; Score 243; DB 1; Length 361;  
 Best Local Similarity 66.7%; Pred. No. 2,2e-21;  
 Matches 46; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSGRGCGTSDNAAAALAAQIRKIPKRYLFOFHNLSNL 60  
 DB 110 TVMGQFQIDRSGRGCGTSDNAAAALAAQIRKIPKRYLFOFHNLSNL 169  
 QY 61 DRIEFKPTPE 69  
 DB 170 DRIEFKAPPE 178

RESULT 3  
 ID VAL1 CLVJ STANDARD; PRT: 358 AA.  
 AC P14982;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DE ALL PROTEIN (40.4 KDA protein update)  
 OS Cassava latent virus (strain West Konyak 844).  
 OS Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 RN PR00227; GEMCOATCLVLI.  
 RN SEQUENCE FROM N.A.  
 RA Stanley J., Gay M.R.;  
 RT Nucleotide sequence of cassava latent virus DNA. \*;  
 RL Nature 301:460-462(1983).  
 CC 1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; X17095; CAA34953.1; -  
 DR PIR; S07594; S07594.  
 DR INTERPRO; IPR001301; -  
 DR PFAM; PF00759; GeminLALL; 1.  
 DR PRINTS; PR00227; GEMCOATL1.  
 DR PRINTS; PR00228; GEMCOATCLVLI.  
 DR NP-BINDing. 227  
 FT SEQUENCE 358 AA; 40435 MW; 1DB16HBCB2D5E2C CRC64;  
 SQ

Query Match 64.2%; Score 230; DB 1; Length 358;  
 Best Local Similarity 60.0%; Pred. No. 7,3e-20;  
 Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSGRGCGTSDNAAAALAAQIRKIPKRYLFOFHNLSNL 60  
 DB 109 TVMGQFQIDRSGRGCGTSDNAAAALAAQIRKIPKRYLFOFHNLSNL 168

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; J02057; -; NOT\_ANNOTATED\_CDS.  
 DR INTERPRO; IPR001301; -  
 DR PFAM; PF00759; GeminLALL; 1.  
 DR PRINTS; PR00227; GEMCOATL1.  
 DR PRINTS; PR00228; GEMCOATCLVLI.  
 DR NP-BINDing. 220  
 FT NP-BIND 220  
 FT SEQUENCE 358 AA; 40346 MW; ED13E753EE92D69 CRC64;  
 SQ

Query Match 64.2%; Score 230; DB 1; Length 358;  
 Best Local Similarity 60.0%; Pred. No. 7,3e-20;  
 Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSGRGCGTSDNAAAALAAQIRKIPKRYLFOFHNLSNL 60  
 DB 109 TVMGQFQIDRSGRGCGTSDNAAAALAAQIRKIPKRYLFOFHNLSNL 168  
 QY 61 DRIEFKPTPE 70  
 DB 169 DRIEFQEPAP 178

RESULT 4  
 ID VAL1 CLVJ STANDARD; PRT: 358 AA.  
 AC P14972;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DE ALL PROTEIN (40.4 KDA protein update)  
 OS Cassava latent virus (strain Nigerian).  
 OS Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 RN PR00227; GEMCOATCLVLI.  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE-90174930; PubMed-2308831;  
 RA Morris B., Coates L., Love S., Richardson K., Eddy P.;  
 RT Nucleotide sequence of the infectious cloned DNA components of  
 RL African cassava mosaic virus (strain Nigerian).  
 CC 1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; X17095; CAA34953.1; -  
 DR PIR; S07594; S07594.  
 DR INTERPRO; IPR001301; -  
 DR PFAM; PF00759; GeminLALL; 1.  
 DR PRINTS; PR00227; GEMCOATL1.  
 DR PRINTS; PR00228; GEMCOATCLVLI.  
 DR NP-BINDing. 227  
 FT SEQUENCE 358 AA; 40435 MW; 1DB16HBCB2D5E2C CRC64;  
 SQ

Query Match 64.2%; Score 230; DB 1; Length 358;  
 Best Local Similarity 60.0%; Pred. No. 7,3e-20;  
 Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSGRGCGTSDNAAAALAAQIRKIPKRYLFOFHNLSNL 60  
 DB 109 TVMGQFQIDRSGRGCGTSDNAAAALAAQIRKIPKRYLFOFHNLSNL 168

```

Qy 61 DRFDKTEP 70
Db 169 DRFDKTEP 178

RESULT 5
VALL_TYLA
ID VALL_TYLA STANDARD: PRT: 362 AA.
AC P36279:1094 (Rel. 29, Created)
DT 01-JUN-1991 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN (C1 PROTEIN).
CN C1.
OS Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RX MEDLINE=91020984; PubMed=8423446;
RA DRY I.B., Riden J.E., Krake L.B., Mullineux P.M., Kezalan M.A.:
RT "Nucleotide sequence and genome organization of tomato leaf curl
geminiavirus.";
RL J. Gen. Virol. 74:147-151(1993).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL1 PROTEIN FAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on
its use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL: X51983; -: NOT_ANNOTATED_CDS.
DR PIR: A36214; OQVMI.
DR INTERPRO: IPR001191; -.
DR PIR: PF00799; Gemini_ALL1.
DR PRINTS: PR00227; GEMCOATALL.
DR PRINTS: PR00228; GEMCOATCLV1.
KW ATP-binding, 221 228 ATP (POTENTIAL).
FT NP-BIND 221 228
SQ SEQUENCE 362 AA: 41197 MW: 343E7184B704098 CRC64;

Query Match 62.0%; Score 222; DB 1; Length 362;
Best local Similarity 52.9%; Pred. No. 6.5e-19;
Matches 4; Conservative 5; Mismatches 15; Indels 16; Gaps 1;
Qy 1 TLWGEFQVDSRGCGQTSNDAAALAAASSAAALQITREKIPKYLQPHNLNSL 60
Db 110 TLENGEFOIDGRSARGGQSDANDVAQALWTGSKSEALNVLRELAPKDYVLQPHNLNSL 169

Qy 61 DRI-
Db 170 DRIFFLEVPYSPFSSSFORVPE 194

RESULT 6
VALL_ABNVW
ID VALL_ABNVW STANDARD: PRT: 355 AA.
AC P21947:1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN.
CN AC1
OS Abutilon mosaic virus (isolate West India).
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RX MEDLINE=91020984; PubMed=2219703;
RA Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,
Rivera-Bustamante R.F.:
RT "Complete nucleotide sequence of pepper huasteco virus: analysis and
comparison with bipartite geminiviruses.";
RL J. Gen. Virol. 74:2235-2231(1993).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL1 PROTEIN FAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on
its use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL: X70418; CAA49856.1; -.
DR PIR: S31875; S31875.
DR PIR: JQ2300; JQ2300.
DR INTERPRO: IPR001191; -.
DR INTERPRO: IPR001301; -.

```

DR PFAM: PF00799; Gemin1\_ALI. 1.  
 DR PRINTS: PR00227; GEMCOATLVL.  
 KA ATP: 21-Jun-1994 (Rel. 29, Last annotation update)  
 FT NP\_BIND 221 228 ATP (BY SIMILARITY).  
 SQ SEQUENCE 349 AA: 39722 MW: 53475 D5F4E75C56370F4 CRC64;

Query Match 60.3%; Score 216; DB 1; Length 349;  
 Best Local Similarity 58.0%; Pred. No. 3.3e-18;  
 Matches 41; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLVGEFQVDSRGSGCOTNDAAALAAALQIIRKIPKYLFOFHNSNL 60  
 DB 110 TWMEQFQVDSRGSGCOTNDAAALAAALQIIRKIPKYLFOFHNSNL 169

QY 61 DRIFDKTPPEP 70

DB 170 NRIFQVPEP 179

# RESULT 8

VALI\_TYLCU STANDARD: PRT: 359 AA.

ID VALI\_TYLCU  
 AC P38609:  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DE AL1 PROTEIN (C1 PROTEIN).  
 GN C1.

OS Tomato yellow leaf curl virus (strain Murcia) (TYLCV).  
 OS Viruses; ssDNA viruses; Geminiviridae; Begomovirus.

RN [1] JENSEN FROM N.A.

RX MEDLINE:9425636; PubMed-8198442;

RA Norris E., Hidalgo E., Accotto G.,

RT "High similarity among the tomato yellow leaf curl virus isolates

from the west Mediterranean basin: the nucleotide sequence of an

infectious clone from Spain.\*;

Arch. Virol. 135:165-170(1994).

CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcements>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: Z57511; CRA81026.1.

DR PIR: S39211; S39211.

DR INTERPRO: IPR001191.

DR PFAM: PF00799; Gemin1\_ALI. 1.

DR PRINTS: PR00227; GEMCOATLVL.

KA ATP: 21-Jun-1994 (Rel. 29, Last annotation update)

FT NP\_BIND 221 228 ATP (POTENTIAL).

SQ SEQUENCE 359 AA: 41065 MW: 20170A51EF80A3EC CRC64;

Query Match 60.3%; Score 216; DB 1; Length 359;  
 Best Local Similarity 58.0%; Pred. No. 3.3e-18;  
 Matches 40; Conservative 9; Mismatches 20; Indels 0; Gaps 0;

QY 2 LVWGEFQVDSRGSGCOTNDAAALAAALQIIRKIPKYLFOFHNSNL 61

DB 111 LMGTFQVDSRGSGCOTNDAAALAAALQIIRKIPKYLFOFHNSNL 170

QY 62 RIFDKTPPEP 70

DB 171 RYFQVPEP 179

# RESULT 9

VALI\_TMOV STANDARD: PRT: 353 AA.

ID VALI\_TMOV  
 AC P05175:  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE AL1 PROTEIN (40.2 KDA PROTEIN).

GN bean golden mosaic virus.

OS Viruses; ssDNA viruses; Geminiviridae; Begomovirus.

RN [1] JENSEN FROM N.A.

RA Howarth A.J., Caton J., Bossert M., Goodman R.M.;

RT "Nucleotide sequence of bean golden mosaic virus and a model for gene

proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).

CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcements>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: M10070; AAA46318.1.

DR INTERPRO: IPR001191.

DR PFAM: PF00799; Gemin1\_ALI. 1.

DR PRINTS: PR00227; GEMCOATLVL.

KA ATP: 21-Jun-1994 (Rel. 29, Last annotation update)

FT NP\_BIND 222 229 ATP (POTENTIAL).

SQ SEQUENCE 353 AA: 40190 MW: 80FA775DF629A34 CRC64;

# Query Match

Best Local Similarity 60.1%; Score 215; DB 1; Length 353;  
 Matches 41; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLVGEFQVDSRGSGCOTNDAAALAAALQIIRKIPKYLFOFHNSNL 60

DB 110 TWMEQFQVDSRGSGCOTNDAAALAAALQIIRKIPKYLFOFHNSNL 169

QY 61 DRIFDKTPPEP 70

DB 170 ERIFRVPPEP 179

# RESULT 10

VALI\_TMOV STANDARD: PRT: 361 AA.

ID VALI\_TMOV  
 AC Q06657: 1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE AL1 PROTEIN.

GN Tomato mottle virus (isolated from Florida) (TMov).

OS Viruses; ssDNA viruses; Geminiviridae; Begomovirus.

RN [1] JENSEN FROM N.A.

RX MEDLINE-9310768; PubMed-1469361;

RA Abouid A.M., Polston J.E., Hiebert E.;

RT "The nucleotide sequence of tomato mottle virus, a new geminivirus

isolated from tomatoes in Florida.\*;

J. Gen. Virol. 73:3225-3229(1992).

CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use, and this statement is as content in no way  
 CC modified and this statement is not removed usage by commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L14460; AAC32414.1; -;  
 DR TRIPREP; 001870;  
 DR INTERPRO: IPRO01301; -;  
 DR PFM: PF00799; Gemini: ALL: 1;  
 DR PRINTS: PR00227; GEMCOATALL;  
 DR PRINTS: PR00228; GEMCOATCLVL1;  
 DR KW ATP-binding. 222 229 ATP (BY SIMILARITY).  
 DR NP-BIND 358 AA; 40516 MW; 9138B65CEAC6950 CRC64;  
 DR SEQUENCE 358 AA; 40516 MW; 9138B65CEAC6950 CRC64;  
 FT  
 SQ  
 Query Match 59.8%; Score 214; DB 1; Length 361;  
 Best Local Similarity 55.7%; Pred. No. 5.7e-18;  
 Matches 39; Conservative 15; Mismatches 16; Indels 0; Gaps 0;  
 Oy 1 TLWGSEFVQDGRSGCQTSNDAAALAAASSAAALQIIREKIPKYLEFQPHNLSNL 60  
 Db 110 TLWGSEFVQDGRSGCQTSNDAAALAAASSAAALQIIREKIPKYLEFQPHNLSNL 159  
 Oy 61 DRIFDKTPPEP 70  
 Db 170 ERIFAKAPEP 179  
 RESULT 11  
 VAL1:TYLVC STANDARD; PRT: 359 AA.  
 AC P27260;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE ALL PROTEIN (C1 PROTEIN).  
 OS Tomato yellow leaf curl virus (strain Marmande) (TYLVCV).  
 OS Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=92107660; PubMed=1840676;  
 RA Krensenborn B.; Bendahmane M.; Matzeit V.; Accotto G.P.; Crespi S.;  
 RA Grunewald B.; leaf curl virus from Sardinia is a  
 RA tomato yellow leaf curl virus.  
 EL Nucleic Acids Res. 19,6763-6769(1991)  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use, and this statement is as content in no way  
 CC modified and this statement is not removed usage by commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X61153; CAA43466.1; -;  
 DR TRIPREP; 522593;  
 DR INTERPRO: IPRO01301; -;  
 DR PFM: PF00799; Gemini: ALL: 1;  
 DR PRINTS: PR00227; GEMCOATALL;  
 DR PRINTS: PR00228; GEMCOATCLVL1;  
 DR KW ATP-binding. 220 227 ATP (POTENTIAL).  
 DR NP-BIND 359 AA; 40733 MW; 9717B4A07C93EAF7 CRC64;  
 DR SEQUENCE 359 AA; 40733 MW; 9717B4A07C93EAF7 CRC64;  
 FT  
 SQ

Query Match 59.2%; Score 212; DB 1; Length 359;  
 Best Local Similarity 56.5%; Pred. No. 9.8e-18;  
 Matches 39; Conservative 10; Mismatches 20; Indels 0; Gaps 0;  
 Oy 2 TLWGSEFVQDGRSGCQTSNDAAALAAASSAAALQIIREKIPKYLEFQPHNLSNL 61  
 Db 111 LKWTUFDGRSGCQQTANDATAYAKANGSKQALQIKELAPROYLVLFHFNLSNL 170  
 Oy 62 RIFUKTPPEP 70  
 Db 171 KYEQVPPAP 179  
 RESULT 12  
 VAL1:BCTV STANDARD; PRT: 358 AA.  
 AC P14991;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-AUG-1992 (Rel. 23, Last annotation update)  
 DE ALL PROTEIN (40.8 KDA PROTEIN).  
 OS Beet curly top virus (BCTV).  
 OS Viruses: ssDNA viruses; Geminiviridae; Curtovirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Stanley J.; Markham P.G.; Callis R.J.; Pinner M.S.;  
 RA The nucleotide sequence of an infectious clone of the geminivirus  
 RA beet curly top virus.  
 EL EMBO J. 5:1761-1767(1986).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use, and this statement is as content in no way  
 CC modified and this statement is not removed usage by commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X04144; -; NOT\_ANNOTATED\_CDS.  
 DR INTERPRO: IPRO01191; -;  
 DR PFM: PF00799; Gemini: ALL: 1;  
 DR PRINTS: PR00227; GEMCOATALL;  
 DR PRINTS: PR00228; GEMCOATCLVL1;  
 DR KW ATP-binding. 222 229 ATP (POTENTIAL).  
 DR NP-BIND 358 AA; 40889 MW; 39A45FE3C0B9C333 CRC64;  
 DR SEQUENCE 358 AA; 40889 MW; 39A45FE3C0B9C333 CRC64;  
 FT  
 SQ  
 Query Match 58.1%; Score 208; DB 1; Length 358;  
 Best Local Similarity 55.7%; Pred. No. 2.9e-17;  
 Matches 39; Conservative 13; Mismatches 18; Indels 0; Gaps 0;  
 Oy 1 TLWGSEFVQDGRSGCQTSNDAAALAAASSAAALQIIREKIPKYLEFQPHNLSNL 60  
 Db 110 TLWGSEFVQDGRSGCQQTANDATAYAKANGSKQALQIKELAPROYLVLFHFNLSNL 169  
 Oy 61 DRIFDKTPPEP 70  
 Db 170 QKIFORPPDP 179  
 RESULT 13  
 VAL1:TYLVC STANDARD; PRT: 357 AA.  
 AC P27259;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE ALL PROTEIN (C1 PROTEIN).  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=92107660; PubMed=1840676;  
 RA Krensenborn B.; Bendahmane M.; Matzeit V.; Accotto G.P.; Crespi S.;  
 RA Grunewald B.; leaf curl virus from Sardinia is a  
 RA tomato yellow leaf curl virus.  
 EL Nucleic Acids Res. 19,6763-6769(1991)  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use, and this statement is as content in no way  
 CC modified and this statement is not removed usage by commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X61153; CAA43466.1; -;  
 DR TRIPREP; 522593;  
 DR INTERPRO: IPRO01301; -;  
 DR PFM: PF00799; Gemini: ALL: 1;  
 DR PRINTS: PR00227; GEMCOATALL;  
 DR PRINTS: PR00228; GEMCOATCLVL1;  
 DR KW ATP-binding. 220 227 ATP (POTENTIAL).  
 DR NP-BIND 359 AA; 40733 MW; 9717B4A07C93EAF7 CRC64;  
 DR SEQUENCE 359 AA; 40733 MW; 9717B4A07C93EAF7 CRC64;  
 FT  
 SQ

```

OS Tomato yellow leaf curl virus (TYLCV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN
RN [1] SEQUENCE FROM N.A.
RX MEDLINE-92024070; PubMed-1926771;
RA Navot N., Pichersky E., Zeldan M., Zamir D., Czosnek H.:
RT "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
RL Virology 185:151-161(1991).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL; X15656; CA33568.1;
DR PIR; D40779; QOCVCI
DR INTERPRO: IPR001191;
DR PFAM: PF00799; Gemini_ALI: 1.
DR PRINTS: PR00227; GEMCOATALL.
DR PRINTS; PR00228; GEMCOATCLVL.
KW ATP-binding; 319 226 ATP (POTENTIAL).
FT NP_BIND 319 226
SQ SEQUENCE 357 AA; 40678 MW; 939A9668EA3B2A7 CRC64;

Query Match 53.1%; Score 190; DB 1; Length 357;
Best Local Similarity 60.7%; Pred. No. 3.8e-15;
Matches 37; Conservative 11; Mismatches 13; Indels 0; Gaps 0;

QY 4 WGFQVDSRAGCGQTSNDAAALAAASSAAALQIREKIPERYLQPHNLSNDRIF 63
DB 111 FGVSQIDNSRAGCGQSDNDAAALAAASSAAALQIREKIPERYLQPHNLSNDRIF 170
QY 64 F 64
DB 171 F 171

RESULT 14
VALI_SICV STANDARD: PRT: 347 AA.
AC P29048; 1992 (Rel. 24; Created)
DT 01-DEC-1992 (Rel. 24; Last sequence update)
DT 01-DEC-1992 (Rel. 24; Last annotation update)
DE ALL PROTEIN.
OS Squash leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE-91082449; PubMed-1994668;
RA "Zhang et al. S.G. Lazdins J.B. The amino acid sequence of the cloned genomic
RT components of a bipartite squash leaf curl geminivirus with a broad
RT host range phenotype.";
RL Virology 180:58-69(1991).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL; M38183; AAC32410.1; ALT_INIT.

```

```

DR PIR; C36785; QOCVSI.
DR INTERPRO: IPR001191;
DR PFAM: PF00799; Gemini_ALI: 1.
DR PRINTS: PR00227; GEMCOATALL.
DR PRINTS; PR00228; GEMCOATCLVL.
KW ATP-binding; 218 225 ATP (POTENTIAL).
FT NP_BIND 218 225
SQ SEQUENCE 347 AA; 42493 MW; AFDABDD122110R CRC64;

Query Match 32.1%; Score 115; DB 1; Length 347;
Best Local Similarity 30.4%; Pred. No. 2.6e-06;
Matches 24; Conservative 14; Mismatches 24; Indels 1; Gaps 1;

QY 5 GFQVDSRAGCGQTSNDAAALAAASSAAALQIREKIPERYLQPHNLSNDRIF 64
DB 116 GQYKVSQ---GSKSNKQDVYHNAVNGSEALDIKAGDKFTFVYHNLNANVERLF 171
QY 65 QKPEEP 70
DB 172 QKPEEP 177

RESULT 15
TFPA_DROME STANDARD: PRT: 577 AA.
ID TFPA_DROME
DT 01-FEB-1994 (Rel. 28; Created)
DT 01-NOV-1995 (Rel. 32; Last sequence update)
DT 01-OCT-1996 (Rel. 34; Last annotation update)
DE TRANSCRIPTION INITIATION FACTOR IIF, ALPHA SUBUNIT (TFIIF-ALPHA)
DE DE (TRANSCRIPTION FACTOR 5, LARGE CHAIN) (TF5A).
GN TFIIF-ALPHA.
OS Drosophila melanogaster (Fruit fly).
OC Insecta; Arthropoda; Insecta; Hexapoda; Insecta;
OC Eukaryota; Metazoa; Echinodermata; Chordata; Brachyopoda; Mammalia;
OC Ephydroidea; Drosophilidae; Drosophila.
RN
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE-93219133; PubMed-8464745;
RA Gong D.-W., Horikoshi M., Nakatani Y.;
RT Analysis of cDNA encoding Drosophila transcription initiation factor
RT TFIIF (TFIIF) (1994).
RL Nucleic Acids Res. 21:1492-1492(1993).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE-93219101; PubMed-8464716;
RA Kephart D., Price M.P., Burton Z.F., Finkelstein A.,
RA Grenblatt J., Price D.H.; DNA with sequence similarity to human
RT transcription factor RAP74.
RL Nucleic Acids Res. 21:1319-1319(1993).
RN [3]
RN FUNCTION.
RX MEDLINE-94230464; PubMed-8175788;
RA Kephart D., Wang B.O., Burton Z.F., Price D.H.;
RT Functional analysis of Drosophila factor 5 (TFIIF), a general
RT transcription factor (1994).
RL J. Biol. Chem. 269:13536-13543(1994).
CC -1- FUNCTION: TFIIF IS A GENERAL TRANSCRIPTION INITIATION FACTOR THAT
CC BINDS TO RNA POLYMERASE II AND HELPS TO RECRUIT IT TO THE
CC INITIATION COMPLEX IN COLLABORATION WITH TFIIB. IT PROMOTES
CC TRANSCRIPTION ELONGATION.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- SIMILARITY: TO OTHER TRANSCRIPTION FACTOR IIF, ALPHA SUBUNIT.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X60922; CANA3391.1; .  
 DR EMBL: X60922; CANA3391.1; .  
 DR PTR: S30237; S30237.

DR TRANSFAC: T02173; -.  
 DR FLYBASE: Fgn010282; TFIIF-alpha.

KW Transcription regulation: DNA-binding; Nuclear protein.  
 FT CONFLICT 20 20 R->A (IN REF. 2).

FT CONFLICT 76 76 G->E (IN REF. 2).  
 FT CONFLICT 576 526 A->G (IN REF. 3).

SO SEQUENCE 577 AA: 64453 MW: 862667397F505B3 CRC64;

Query Match 19.3%; Score 69; DB 1; Length 577;

Best Local Similarity 36.4%; Pred. No. 1.3;

Matches 20; Conservative 9; Mismatches 20; Indels 6; Gaps 1;

Q7 13 SARGCCGTSNDAAALANSSAMALQITR-----KIDPKYLFQPHLNSHLD 61

Db 10 SAAGSGSTSAARAASVASGSSASSANVQEFIRVPKPKKHWHWFENATLAVD 54

Search completed: February 3, 2001, 02:24:25

Job time: 637 sec

---

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:17:37 : Search time 118.74 seconds  
(without alignments)  
40.029 Million cell updates/sec

Title: US-09-289-346a-3

Percent score: 359

Sequence: 1 TLVWGFQVQDSARGCQT.....FQFHNLNSLDLRFDTKTPP 70

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PTR.66:\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than the actual score. The score is being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	332	92.7	352	1 QOCVLT	Al1 protein - toma
2	243	97.9	361	1 QOCVLT	Al1 protein - toma
3	230	64.2	358	2 S07594	hypothetical prote
4	222	62.0	362	1 Q1R887	Al1 protein - toma
5	217	60.3	355	1 QOCVM1	AV1 protein - abut
6	216	60.3	349	2 J02400	replicase - pepper
7	216	60.3	349	2 S34875	Al1 protein - pepp
8	214	59.8	358	1 J01870	Al1 protein - toma
9	214	59.8	358	1 J01870	Al1 protein - toma
10	213	59.5	351	2 J02327	Al1 protein - Ind
11	212	59.2	359	2 S22593	hypothetical prote
12	208	58.1	385	2 S28360	Al1 protein - beet
13	207	57.8	360	2 S59885	replication-associ
14	203	56.7	359	2 S39235	gene Cl protein -
15	203	56.7	359	2 S39235	Al1 protein - toma
16	120	33.5	131	2 S45059	Al1 protein (clone
17	115	32.1	347	1 QOCV51	Al1 protein - squa
18	69	19.3	577	2 S30237	transcription init
19	66	18.4	508	2 S59870	fork head domain p
20	66	18.4	540	2 T36175	probable large ATP
21	65	18.2	671	2 A35912	homeotic protein o
22	63	17.6	352	2 T44362	acetyl-CoA C-acety
23	62	17.5	352	2 T44362	hypothetical prote
24	62	17.5	1077	2 A44967	hypothetical prote
25	62	17.3	409	2 T16781	hypothetical prote
26	61.5	17.2	337	2 F83408	hypothetical prote
27	61.5	17.2	705	2 T31157	hypothetical prote
28	61.5	17.2	1541	2 T02831	AAA protein L4171
29	61	17.0	339	2 T26328	hypothetical prote

30 61 17.0 371 2 B69502 alcohol dehydrogen  
31 61 17.0 1265 2 T47626 structural mainten  
32 61 17.0 1510 2 T33100 hypothetical prote  
33 60.5 16.9 509 2 T02942 O-succinylhomoseri  
34 60.5 16.9 883 2 S04722 ecdysone-induced p  
35 60.5 16.9 883 2 S04722 74E  
36 60 16.8 314 2 J05273 paired type homeob  
37 59.5 16.6 201 2 A81380 probable flagellar  
38 59.5 16.6 470 2 T46814 gamma-aminobutyrat  
39 59.5 16.6 481 2 T07091 probable phosphos  
40 59.5 16.6 509 2 T02940 O-succinylhomoseri  
41 59 16.5 380 2 S70594 ubiquinol--cytochr  
42 59 16.5 458 2 T47130 calcium recepto  
43 59 16.5 458 2 T47130 5-aminolevulinate  
44 59 16.5 513 1 S2CHLE  
45 59 16.5 3551 1 S21010 filamentous hemagg

#### ALIGNMENTS

RESULT 1

QOCVLT

Al1 protein - tomato golden mosaic virus

C:Species: Tomato golden mosaic virus

A:Note: host Nicotiana sp. (tobacco)

C:Date: 28-Aug-1985 #sequence\_revision 28-Aug-1985 #text\_change 08-Apr-1994

C:Accession: A04170 Steirn, V.E.; Coutts, R.H.A.; Buck, K.W.

EMBO J. 3, 2197-2205, 1984

A:Title: Complete nucleotide sequence of the infectious cloned DNA components of toma

A:Reference number: A04163

A:Accession: A04170

A:Molecule type: DNA

A:Residues: 1-352 <HAMP>

C:Comment: The genome consists of two circular, single-stranded DNA components, DNA A

C:Genetics:

A:Map position: segment A

C:Superfamily: tomato golden mosaic virus Al1 protein

Query Match Similarity 92.7%; Score 332; DB 1; Length 352;

Best Local 92.9%; Pred: No. 1-1e-30;

Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 1 TLVWGFQVQDSARGCQTSDNDAALAAASSAAALQIIRK1PKYKLFQFHNLNSL 60

Db 111 TLVWGFQVQDSARGCQTSDNDAALAAASSAAALQIIRK1PKYKLFQFHNLNSL 170

Oy 61 DRIFDKTTPP 70

Db 171 DRIFDKTTPP 180

RESULT 2

QOCVPT

Al1 protein - potato yellow mosaic virus (isolate Venezuela)

C:Species: potato yellow mosaic virus

C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Jun-2000

C:Accession: J00364

R:Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.

J. Gen. Virol. 72, 1515-1520, 1991

A:Title: The nucleotide sequence of the infectious cloned DNA components of potato ye

A:Reference number: J00362; MUID:91311403

A:Accession: J00364

A:Molecule type: not shown

A:Residues: 1-361 <COU>

C:Cross-references: GB:D00940; NID:g222458; PIDN:BAA00782.1; PID:g222459

C:Genetics:

A:Map position: segment A

C:Superfamily: tomato golden mosaic virus Al1 protein

```

Query Match      67.9%; Score 243; DB 1; Length 361;
Best Local Similarity 66.7%; Pred. No. 1,9e-21;
Matches 46; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 1 TLVKGFEQVDRSGARGGQTSNDAAAALQITREKIPKYLQFQHNLSNL 60
DB 110 TLMWGEFQIDGRSGARGGQTSNDAAAALQITREKIPKYLQFQHNLSNL 169

QY 61 DRIFQKTPPE 69
DB 170 DRIFQKTPPE 178

RESULT 3
S07594
hypothetical protein, 40.4k - cassava latent virus (Nigerian isolate)
C:Species: cassava latent virus
C:Accession: S07594
C:Accession: S07594 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
R:Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.
Nucleic Acids Res. 18, 197-198, 1990
A>Title: Nucleotide sequence of the infectious cloned DNA components of African cassava
A:Reference number: S07590; MUID:90174930
A:Accession: S07594
A:Status: translation not shown
A:Residues: 1-358 <OR>
A:Cross-References: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376
C:Genetics:
A:Map position: segment DNA1
C:Superfamily: tomato golden mosaic virus ALL protein

Query Match      64.3%; Score 230; DB 2; Length 356;
Best Local Similarity 60.0%; Pred. No. 6.1e-19;
Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

QY 1 TLVKGFEQVDRSGARGGQTSNDAAAALQITREKIPKYLQFQHNLSNL 60
DB 109 TVMGQFQIDGRSGARGGQTSNDAYAKALNGSKSEALNVRELQPKDYVLOFHNLSNL 168

QY 61 DRIFQKTPPE 70
DB 169 DRIFQKTPPE 178

RESULT 4
S01887
ALL protein - tomato yellow leaf curl virus (strain Australia)
N:Alternate names: Cl protein
C:Species: tomato yellow leaf curl virus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C:Accession: J01887
R:Dry, I.B.; Rigden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.
J. Gen. Virol. 74, 147-151, 1993
A>Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.
A:Reference number: J01887; MUID:93139778
A:Accession: J01887
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-362 <OR>
A:Cross-References: GH:S53251
C:Superfamily: tomato golden mosaic virus ALL protein

Query Match      62.0%; Score 222; DB 1; Length 362;
Best Local Similarity 52.9%; Pred. No. 5.1e-18;
Matches 45; Conservative 9; Mismatches 15; Indels 16; Gaps 1;

QY 1 TLVKGFEQVDRSGARGGQTSNDAAAALQITREKIPKYLQFQHNLSNL 60
DB 110 TLMWGEFQIDGRSGARGGQTSNDAAAALQITREKIPKYLQFQHNLSNL 169

```

```

DB 110 TLMWGEFQIDGRSGARGGQTSNDAYAKALNTGSKSEALNVRELQPKDYVLOFHNLSNL 169

QY 61 DRIFQKTPPE 69
DB 170 DRIFQKTPPE 178

RESULT 5
Q0CVW1
AVI protein - abutilon mosaic virus (isolate West India)
C:Species: abutilon mosaic virus
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Apr-1994
R:Prischmann, A.; Ziegler, G.; Jeske, H.
Virology 178, 461-468, 1990
A>Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as
A:Reference number: A36214; MUID:91020984
A:Accession: A36214
A:Molecule type: DNA
A:Residues: 1-355 <PR>
A:Cross-References: EMBL:X15983
C:Genetics:
A:Map position: segment A
C:Superfamily: tomato golden mosaic virus ALL protein

Query Match      60.6%; Score 217; DB 1; Length 355;
Best Local Similarity 58.6%; Pred. No. 1.9e-17;
Matches 41; Conservative 11; Mismatches 18; Indels 0; Gaps 0;

QY 1 TLVKGFEQVDRSGARGGQTSNDAAAALQITREKIPKYLQFQHNLSNL 60
DB 110 TLMWGEFQIDGRSGARGGQTSNDAYAKALNGDVQSALNLKEEDQPKDYVLOHNIRSNL 169

QY 61 DRIFQKTPPE 70
DB 170 DRIFQKTPPE 179

RESULT 6
J02300
replicase - pepper huasteco virus (component A)
N:Alternate names: coat ALL protein
C:Species: pepper huasteco virus
C:Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Sep-1999
C:Accession: J02300
R:Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante,
J. Gen. Virol. 74, 2225-2231, 1993
A>Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b
A:Reference number: J02299; MUID:94015007
A:Accession: J02300
A:Residues: 1-349 <OR>
A:Cross-References: GH:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
C:Superfamily: tomato golden mosaic virus ALL protein

Query Match      60.3%; Score 216; DB 2; Length 349;
Best Local Similarity 58.6%; Pred. No. 2.4e-17;
Matches 41; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLVKGFEQVDRSGARGGQTSNDAAAALQITREKIPKYLQFQHNLSNL 60
DB 110 TVMGQFQIDGRSGARGGQTSNDAYAKALNSAEALQITKEEDQPKDYVLOFHNLSNA 169

QY 61 DRIFQKTPPE 70
DB 170 DRIFQKTPPE 179

RESULT 7
S31875
ALL protein - pepper rizado amarillo virus

```

C:Species: pepper rizado amarillo virus  
 C:Date: 22-Nov-1993 #sequence\_revision 26-May-1995 #text\_change 20-Sep-1999  
 C:Accession: S31875  
 A:Residues: 1-349 <ORF>  
 A:Map position: segment A  
 C:Superfamily: Complete nucleotide sequence of pepper huasteco virus: analysis and comparison with other members of the genus Tospovirus  
 A:Reference number: S31872  
 A:Accession: S31875  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-349 <ORF>  
 A:Cross-references: EMBL:X70418; NID:961023; PIDM:CAA49856.1; PID:961025  
 A:Note: The nucleotide sequence was submitted to the EMBL Data Library, August 1991  
 C:Superfamily: tomato golden mosaic virus A11 protein

Query Match 60.3%; Score 216; DB 2; Length 349;  
 Best Local Similarity 58.6%; Pred. No. 2.4e-17;  
 Matches 41; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

Qy 1 TLWGEFQVGRSANGCQTSNDAAALAAASAAALQIIRKIPKYLFOFHNLNSL 60  
 Db 110 TLWGEFQVGRSANGCQTSNDAAALAAASAAALQIIRKIPKYLFOFHNLNSL 169

Qy 61 DRIFDKTPEP 70  
 Db 170 NRIFQPPPEP 179

## RESULT 8

gene C1 protein - tomato yellow leaf curl virus  
 C:Species: tomato yellow leaf curl virus  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 20-Sep-1999  
 C:Accession: S39211  
 A:Residues: 1-359 <ORF>  
 A:Map position: segment A  
 C:Superfamily: High similarity among the tomato yellow leaf curl virus isolates from the Americas  
 A:Reference number: S39209  
 A:Accession: S39211  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-359 <ORF>  
 A:Cross-references: EMBL:Z25751; NID:9433655; PIDM:CAA81026.1; PID:9433658  
 C:Superfamily: tomato golden mosaic virus A11 protein

Query Match 60.3%; Score 216; DB 2; Length 359;  
 Best Local Similarity 58.0%; Pred. No. 2.5e-17;  
 Matches 40; Conservative 9; Mismatches 20; Indels 0; Gaps 0;

Qy 2 LWKGFQVGRSANGCQTSNDAAALAAASAAALQIIRKIPKYLFOFHNLNSL 61  
 Db 111 LWKGFQVGRSANGCQTSNDAAALAAASAAALQIIRKIPKYLFOFHNLNSL 170

Qy 62 RIFDKTPEP 70  
 Db 171 RVFQVPPAP 179

## RESULT 9

A11 protein - tomato mottle virus (isolate Florida)  
 C:Species: tomato mottle virus  
 C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 07-May-1999  
 C:Accession: J01870  
 A:Residues: 1-329 <ORF>  
 A:Map position: segment A  
 C:Superfamily: tomato mottle virus, a new geminivirus isolated from tomato  
 A:Reference number: J01869; WUID:93107858  
 A:Accession: J01870  
 A:Status: translation not shown  
 A:Molecule type: DNA

A:Residues: 1-358 <ABO>  
 A:Cross-references: GB:L14460  
 C:Genetics:  
 A:Map position: segment A  
 C:Superfamily: tomato golden mosaic virus A11 protein

Query Match 59.8%; Score 214; DB 1; Length 358;  
 Best Local Similarity 55.7%; Pred. No. 4.2e-17;  
 Matches 39; Conservative 15; Mismatches 16; Indels 0; Gaps 0;

Qy 1 TLWGEFQVGRSANGCQTSNDAAALAAASAAALQIIRKIPKYLFOFHNLNSL 60  
 Db 107 TLWGEFQVGRSANGCQTSNDAAALAAASAAALQIIRKIPKYLFOFHNLNSL 166

Qy 61 DRIFDKTPEP 70  
 Db 167 ERIFAKPEP 176

## RESULT 10

A11 protein - Indian cassava mosaic virus  
 N:Alternate names: replication-associated protein  
 C:Species: Indian cassava mosaic virus  
 C:Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 20-Sep-1999  
 C:Accession: J02327; S35883  
 A:Residues: 1-351 <ORF>  
 A:Map position: segment A  
 C:Superfamily: Nucleotide sequence evidence for the occurrence of three distinct whitefly-transmitted viruses  
 A:Reference number: J02326; WUID:94065670  
 A:Accession: J02327  
 A:Molecule type: DNA  
 A:Residues: 1-351 <ORF>  
 A:Cross-references: EMBL:Z24758; NID:9395351; PIDM:CAA80891.1; PID:9584046  
 C:Superfamily: tomato golden mosaic virus A11 protein

Query Match 59.5%; Score 213; DB 2; Length 351;  
 Best Local Similarity 61.2%; Pred. No. 5.4e-17;  
 Matches 41; Conservative 10; Mismatches 16; Indels 0; Gaps 0;

Qy 4 WIEFQVGRSANGCQTSNDAAALAAASAAALQIIRKIPKYLFOFHNLNSL 63  
 Db 113 WIEFQVGRSANGCQTSNDAAALAAASAAALQIIRKIPKYLFOFHNLNSL 172

Qy 64 FDKTPEP 70  
 Db 173 FTKPPEP 179

## RESULT 11

hypothetical protein C4 - tomato yellow leaf curl virus  
 C:Species: tomato yellow leaf curl virus  
 C:Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 20-Sep-1999  
 C:Accession: Z22593  
 A:Residues: 1-359 <ORF>  
 A:Map position: segment A  
 C:Superfamily: tomato golden mosaic virus A11 protein  
 A:Reference number: Z22593  
 A:Accession: Z22593  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-359 <ORF>  
 A:Cross-references: EMBL:X61153; NID:962211; PIDM:CAA43466.1; PID:962217  
 A:Note: The nucleotide sequence was submitted to the EMBL Data Library, August 1991  
 C:Superfamily: tomato golden mosaic virus A11 protein

Query Match 59.2%; Score 212; DB 2; Length 359;  
 Best Local Similarity 56.5%; Pred. No. 7.2e-17;  
 Matches 39; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

QY 2 LWNGEFOVDSRGCGCOTNDAAAALQIIREKIPKYLFOFHLNSNLD 61  
 Db 111 LFWGFOVDSRGCGCOTNDAAAALQIIREKIPKYLFOFHLNSNLD 170  
 QY 62 RIFDKTPEP 70  
 Db 171 KVFQVPAP 179

RESULT 12  
 S:Species: beetle curly top virus  
 A1: protein - beetle curly top virus  
 C:Species: beetle curly top virus  
 C:Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 20-Sep-1999  
 R:Accession: S58360  
 R:Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.  
 E:MO J. 5, 1761-1767, 1986  
 A:Title: The nucleotide sequence of an infectious clone of the geminivirus beetle curly top virus  
 A:Accession: S58360  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-385 <STA>  
 A:Cross-references: GB:M24597; EMBL:X04144; NID:g210678; PIDN:AAA42751.1; PID:g210679  
 C:Superfamily: tomato golden mosaic virus A1 protein

Query Match 58.1%; Score 208; DB 2; Length 385;  
 Best Local Similarity 55.7%; Pred. No. 2,3e-16;  
 Matches 39; Conservative 13; Mismatches 18; Indels 0; Gaps 0;

QY 1 TLWNGEFOVDSRGCGCOTNDAAAALQIIREKIPKYLFOFHLNSNLD 60  
 Db 137 TLWNGEFOVDSRGCGCOTNDAAAALQIIREKIPKYLFOFHLNSNLD 196  
 QY 61 RIFDKTPEP 70  
 Db 197 OKIFORPEP 206

RESULT 13  
 S:Species: replication-associated protein C1 - tomato yellow leaf curl virus  
 C:Species: tomato yellow leaf curl virus  
 C:Date: 14-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 20-Sep-1999  
 R:Accession: S59885  
 R:Hong, Y.; Harrison, B.D.  
 A:Description: Nucleotide sequences from tomato leaf curl viruses from different countries submitted to the EMBL Data Library, February 1995  
 A:Reference number: S58346  
 A:Accession: S59885  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-360 <HON>  
 A:Cross-references: EMBL:248182; NID:g944838; PIDN:CAA98229.1; PID:g974211  
 C:Superfamily: tomato golden mosaic virus A1 protein

Query Match 57.8%; Score 207; DB 2; Length 360;  
 Best Local Similarity 50.6%; Pred. No. 2.7e-16;  
 Matches 40; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

QY 4 WGFVQVDSRGCGCOTNDAAAALQIIREKIPKYLFOFHLNSNLDRI 63  
 Db 113 FGVSVIDGRSGCGCOTNDAAAALQIIREKIPKYLFOFHLNSNLDRI 172

QY 64 PDKTPE 69  
 Db 173 FTFSAE 178

RESULT 14  
 S:Species: tomato yellow leaf curl virus  
 C:Species: tomato yellow leaf curl virus  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 08-Sep-1997  
 R:Accession: S59235  
 R:Crepini, S.; Norris, G.; Vaira, A.; Bosco, D.; Accotto, G.  
 A:Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity.  
 A:Reference number: S59233  
 A:Accession: S59235  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-559 <CRE>  
 A:Cross-references: EMBL:Z28390; NID:g1041671; PID:g1334964  
 C:Superfamily: tomato golden mosaic virus A1 protein

Query Match 56.7%; Score 203; DB 2; Length 359;  
 Best Local Similarity 53.6%; Pred. No. 7.9e-16;  
 Matches 37; Conservative 12; Mismatches 20; Indels 0; Gaps 0;

QY 2 LWNGEFOVDSRGCGCOTNDAAAALQIIREKIPKYLFOFHLNSNLD 61  
 Db 111 LFWGFOVDSRGCGCOTNDAAAALQIIREKIPKYLFOFHLNSNLD 170  
 QY 52 RIFDKTPEP 70  
 Db 171 KVFQVPAP 179

RESULT 15  
 S:Species: tomato yellow leaf curl virus  
 C:Species: tomato yellow leaf curl virus  
 C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 23-Jul-1999  
 R:Accession: D40779  
 R:Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.  
 A:Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin replication-associated protein C1  
 A:Reference number: A40779; MUID:92024070  
 A:Accession: D40779  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-357 <NAV>  
 A:Cross-references: GB:X15656; NID:g62204; PIDN:CAA33688.1; PID:g62207  
 C:Superfamily: tomato golden mosaic virus A1 protein

Query Match 53.1%; Score 190; DB 1; Length 357;  
 Best Local Similarity 60.7%; Pred. NO. 2.9e-14;  
 Matches 37; Conservative 11; Mismatches 13; Indels 0; Gaps 0;

QY 4 WGFVQVDSRGCGCOTNDAAAALQIIREKIPKYLFOFHLNSNLDRI 63  
 Db 111 FGVSVIDGRSGCGCOTNDAAAALQIIREKIPKYLFOFHLNSNLDRI 170

QY 64 P 64  
 Db 171 P 171

Search completed: February 3, 2001, 02:17:38  
 Job time: 2210 sec

GenCone version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:19:40 ; Search time 109.09 Seconds  
(without alignments)  
11.523 Million cell updates/sec

Title: US-09-289-346a-3

Perfit score: 398

Sequence: 1 TLVKGQFQDGRSARGSCQT.....FQEHNLNSLWRFQKTEFP 70

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

1: /cgn2\_6/ptodata/1/aa/5A\_COMB.pep.\*

2: /cgn2\_6/ptodata/1/aa/5B\_COMB.pep.\*

3: /cgn2\_6/ptodata/1/aa/6\_COMB.pep.\*

4: /cgn2\_6/ptodata/1/aa/PCBUS\_COMB.pep.\*

5: /cgn2\_6/ptodata/1/aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	212	59.2	359	3	US-08-809-103B-2
2	212	59.2	359	3	US-08-809-103B-4
3	212	59.2	359	3	US-08-809-103B-6
4	212	59.2	359	3	US-08-809-103B-8
5	613	17.6	103	1	US-08-532-142A-13
6	609	16.5	482	1	US-07-792-864B-17
7	59	16.5	482	1	US-07-792-864B-17
8	59	16.5	482	1	US-08-142-439A-7
9	59	16.5	482	2	US-08-869-477-7
10	58.5	16.3	1713	3	US-08-600-982-24
11	58.5	16.3	1713	4	PCT-US94-10261A-24
12	56.5	15.8	664	3	US-09-295-186-17
13	56	15.6	446	2	US-08-672-514D-11
14	54.5	15.2	485	1	US-08-120-827-1
15	54.5	15.2	485	1	US-08-120-827-1
16	54.5	15.2	485	1	US-08-478-675-1
17	54.5	15.2	738	3	US-08-864-038A-3
18	54	15.1	171	2	US-08-609-049A-22
19	54	15.1	587	2	US-08-871-266B-18
20	54	15.1	587	2	US-09-018-864A-18
21	54	15.1	587	3	US-08-871-267B-24
22	54	15.1	1187	2	US-08-808-794-23
23	53.5	14.9	2100	2	US-08-808-794-23
24	53.5	14.9	2100	3	US-08-772-512A-19
25	52.5	14.7	440	3	US-09-100-664A-2
26	52.5	14.7	440	3	US-09-100-664A-3
27	52.5	14.7	440	3	US-09-100-664A-4
28	52.5	14.7	513	3	US-09-100-193-3

Query Match 59.2%; Score 212; DB 3; Length 359;  
Best Local Similarity 56.5%; Pred. No. 1.2e-20;  
Matches 39; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

29 52.5 14.7 1184 2 US-08-918-914-1 Sequence 1, Appli  
30 52.5 14.7 1184 2 US-08-996-083-3 Sequence 3, Appli  
31 52 14.5 391 5 5229279-2 Patent No. 5229279  
32 52 14.5 391 5 522669-2 Patent No. 522669  
33 52 14.5 391 5 522669-2 Patent No. 522669  
34 52 14.5 391 5 522669-2 Patent No. 522669  
35 52 14.5 391 5 522669-2 Patent No. 522669  
36 52 14.5 391 5 522669-2 Patent No. 522669  
37 51.5 14.4 37 2 US-07-945-283-2 Sequence 2, Appli  
38 51.5 14.4 303 3 US-08-117-952-783 Sequence 783, App  
39 51.5 14.4 303 3 US-09-045-973-1 Sequence 1, Appli  
40 51 14.2 439 1 US-08-445-135-2 Sequence 2, Appli  
41 50 14.0 139 1 US-07-782-681A-1 Sequence 1, Appli  
42 50 14.0 139 1 US-07-782-681A-1 Sequence 1, Appli  
43 50 14.0 98 2 US-07-814-220-2 Patent No. 5496550  
44 50 14.0 98 2 US-07-812-421-2 Sequence 2, Appli  
45 50 14.0 126 1 US-08-083-949-1 Sequence 1, Appli

## ALIGNMENTS

RESULT 1  
US-08-809-103B-2  
; Sequence 2, Application US/08809103B  
; GENERAL INFORMATION:  
; APPLICANT: GRENHORN, BRUNO  
; INVENTOR: GRENHORN, BRUNO  
; TITLE OF INVENTION: TRANSGENIC DNA VIRUS RESISTANT  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: YOUNG & THOMPSON  
; STREET: 745 South 23rd Street  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; COUNTRY OF ORIGIN: U.S.A.  
; APPLICATION NUMBER: US/08/809,103B  
; FILING DATE: 17-MAR-1997  
; CLASSIFICATION: 800  
; PRIOR APPLICATION NUMBER: FR 94.11040  
; APPLICATION DATA: NO PCT/FR95/01192  
; FILING DATE: 15-SEP-1995  
; NAME: PATCH, Andrew J.  
; ATTORNEY/AGENT INFORMATION:  
; REGISTRATION NUMBER: 32,925  
; REFERENCE/DOCKET NUMBER: USB94AL CNR TOM  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 532-2297  
; TELEFAX: (703) 532-2297  
; TELETYPE: 248425 EMBON  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 359 amino acids  
; TYPE: amino acid  
; TOPOLOGY: Linear  
; MOLECULE TYPE: protein  
US-08-809-103B-2

```
QY 2 LVWGEFVDRGARGCCTSNDAANAAALAIIRKIPKYLQFHNLSNLD 61
Db 111 LKNGTFQIDGARGCGQTANDAYAKINAGSKQSDLVKELAPRDYVLFHFNLSNLD 170
QY 62 RIFDKTPEP 70
Db 171 KVFQVFPAP 179

RESULT 2
US-08-809-103B-4
: Sequence 4, Application US/08809103B
: Patent No. 6133505
: GENERAL INFORMATION:
: APPLICANT: GROENBORN, Bruno
: TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
: TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: YOUNG & THOMPSON
: STREET: 745 South 23rd Street
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: FILING DATE: 17-MAR-1997
: APPLICATION NUMBER: US/08/809,103B
: PRIORITY INFORMATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: FR 94.11040
: FILING DATE: 15-SEP-1994
: INVENTOR: PATRICK M. NICHOLS
: FILING DATE: 15-SEP-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: PATCH, Andrew J.
: REGISTRATION NUMBER: 32,925
: REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 521-2297
: TELEFAX: (703) 685-0573
: TELEX: 248425 EMBON
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 359 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-809-103B-4
```

```
Query Match 59.2%; Score 212; Db 3; Length 359;
Best Local Similarity 56.5%; Pred. No. 1.2e-20;
Matches 39; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

QY 2 LVWGEFVDRGARGCCTSNDAANAAALAIIRKIPKYLQFHNLSNLD 61
Db 111 LKNGTFQIDGARGCGQTANDAYAKINAGSKQSDLVKELAPRDYVLFHFNLSNLD 170
QY 62 RIFDKTPEP 70
Db 171 KVFQVFPAP 179
```

## RESULT 3

US-08-809-103B-6

```
: Sequence 6, Application US/08809103B
: Patent No. 6133505
: GENERAL INFORMATION:
: APPLICANT: GROENBORN, Bruno
: TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
: TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: YOUNG & THOMPSON
: STREET: 745 South 23rd Street
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: FILING DATE: 17-MAR-1997
: APPLICATION NUMBER: US/08/809,103B
: PRIORITY INFORMATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: FR 94.11040
: FILING DATE: 15-SEP-1994
: INVENTOR: PATRICK M. NICHOLS
: FILING DATE: 15-SEP-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: PATCH, Andrew J.
: REGISTRATION NUMBER: 32,925
: REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 521-2297
: TELEFAX: (703) 685-0573
: TELEX: 248425 EMBON
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 359 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-809-103B-6
```

```
Query Match 59.2%; Score 212; Db 3; Length 359;
Best Local Similarity 56.5%; Pred. No. 1.2e-20;
Matches 39; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

QY 2 LVWGEFVDRGARGCCTSNDAANAAALAIIRKIPKYLQFHNLSNLD 61
Db 111 LKNGTFQIDGARGCGQTANDAYAKINAGSKQSDLVKELAPRDYVLFHFNLSNLD 170
QY 62 RIFDKTPEP 70
Db 171 KVFQVFPAP 179
```

## RESULT 4

US-08-809-103B-8

```
: Sequence 8, Application US/08809103B
: Patent No. 6133505
: GENERAL INFORMATION:
: APPLICANT: GROENBORN, Bruno
: TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
: TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: YOUNG & THOMPSON
: STREET: 745 South 23rd Street
: CITY: Arlington
: STATE: Virginia
```

```

1  COUNTRY: U.S.A.
2  ZIP: 22202
3  COMPUTER READABLE FORM:
4  PRIORITY: 100
5  COMPUTER FILE: IBM PC
6  OPERATING SYSTEM: PC-DOS/MS-DOS
7  SOFTWARE: Patent In Release #1.0, Version #1.30
8  CURRENT APPLICATION DATA:
9  APPLICATION NUMBER: US/08/809,103B
10 FILING DATE: 17-MAR-1997
11 CLASSIFICATION: 800
12 PRIORITY: 100
13 APPLICATION NUMBER: FR 94.11040
14 FILING DATE: 15-SEP-1994
15 PRIOR APPLICATION NUMBER: WO PCT/FR95/01192
16 FILING DATE: 15-SEP-1995
17 ATTORNEY/AGENT INFORMATION:
18 NAME: PATCH, Andrew J, 925
19 STREET: 10000
20 REFERENCE/DOCKET NUMBER: US94A1 CNR TOM
21 TELECOMMUNICATION INFORMATION:
22 TELEPHONE: (703) 521-2297
23 TELEFAX: (703) 685-0573
24 TELEX: 248425 EMBON
25 INFORMATION FOR SEQ ID NO: 8:
26 SEQUENCE CHARACTERISTICS:
27 LENGTH: 103 amino acids
28 TYPE: amino acid
29 TOPOLOGY: linear
30 MOLECULE TYPE: protein
31 US-08-809-103B-8

32 Query Match 59.28; Score 212; DB 3; Length 359;
33 Best Local Similarity 56.34; Pred. No. 0.28; Indels 0;
34 Matches 39; Conservative 10; Mismatches 20;
35
36 Qy 2 LWKQFQVDRSGRGQGTNDAAALAAASAAALQIIRKIPKYLFOFHNLNSLND 61
37 DB 111 LKQFQVDRSGRGQGTNDAAALAAASAAALQIIRKIPKYLFOFHNLNSLND 170
38
39 Qy 62 RIFQKTPDP 70
40 DB 171 KVFQVPPAP 179
41
42 RESULT 5
43 US-08-552-142A-13
44 Sequence 13; Application US/08552142A
45 Patent No. 5912414
46 GENERAL INFORMATION:
47 APPLICANT: NOLINTRAUB, Harold M.
48 APPLICANT: Lee, Jacqueline E.
49 APPLICANT: Tapscott, Stephen J.
50 APPLICANT: Hollenberg, Stanley M.
51 TITLE OF INVENTION: Neurogenic Differentiation (Neurod) Genes
52 TITLE OF INVENTION: and Proteins
53 TITLE OF INVENTION: Correspondence Address:
54 CORRESPONDENCE ADDRESS:
55 ADDRESSSEE: Christensen O'Connor Johnson Kindness PLLC
56 STREET: 1420 Fifth Avenue, Suite 2800
57 CITY: Seattle
58 STATE: WA
59 COUNTRY: USA
60 ZIP: 98101-2347
61 COMPUTER READABLE FORM:
62 MEDIUM TYPE: FLOPPY disk
63 COMPUTER: IBM PC compatible
64 OPERATING SYSTEM: PC-DOS/MS-DOS
65 SOFTWARE: Patent In Release #1.0, Version #1.25
66 CURRENT APPLICATION DATA:
67 APPLICATION NUMBER: US/08/552.142A

```

```

1  FILING DATE: 02-NOV-1995
2  CLASSIFICATION: 514
3  PRIOR APPLICATION DATA:
4  APPLICATION NUMBER: US 08/239,238
5  FILING DATE: 08-MAY-1994
6  PRIOR APPLICATION DATA:
7  APPLICATION NUMBER: WO PCT/US95/05741
8  FILING DATE: 08-MAY-1995
9  ATTORNEY/AGENT INFORMATION:
10 NAME: Broderick, Thomas F.
11 REFERENCE/DOCKET NUMBER: FPCR-1-8933
12 TELECOMMUNICATION INFORMATION:
13 TELEPHONE: 206-682-8100
14 TELEFAX: 206-225-0709
15 INFORMATION FOR SEQ ID NO: 13:
16 SEQUENCE CHARACTERISTICS:
17 LENGTH: 103 amino acids
18 TYPE: amino acid
19 TOPOLOGY: linear
20 MOLECULE TYPE: protein
21 ORIGINAL SOURCE:
22 ORGANISM: Homo sapiens
23 US-08-352-142A-13

24 Query Match 17.68; Score 63; DB 1; Length 103;
25 Best Local Similarity 33.34; Pred. No. 0.28; Indels 1;
26 Matches 19; Conservative 11; Mismatches 25;
27
28 Qy 14 ARGCGTNSDAALAA--SSAAALQIIRKIPKYLFOFHNLNSLDRIFDKTP 68
29 DB 6 SNGTGRAGEAAAGRSUGAAAXAAGRRYKANDRRNRNMNINAALDALRSVLP 62
30
31 RESULT 6
32 US-08-737-524B-27
33 Sequence 27; Application US/08737524B
34 Patent No. 5912414
35 GENERAL INFORMATION:
36 APPLICANT: CARL SAYERIO FALCO
37 APPLICANT: DOMINICK ANTHONY GUIDA, JR.
38 APPLICANT: MARY ELIZABETH HARNETT LOCKE
39 TITLE OF INVENTION: HORMONE RECEPTORS, OLIGOMERIC
40 TITLE OF INVENTION: GENES AND METHODS FOR INCREASING
41 TITLE OF INVENTION: THE METHYLONE CONTENT OF THE SEEDS
42 NUMBER OF SEQUENCES: 27
43 CORRESPONDENCE ADDRESS:
44 ADDRESSSEE: E. I. DU PONT DE NEMOURS AND COMPANY
45 STREET: 1007 MARKET STREET
46 CITY: WILMINGTON
47 STATE: DELAWARE
48 COUNTRY: UNITED STATES OF AMERICA
49 ZIP: 19898
50 COMPUTER READABLE FORM:
51 MEDIUM TYPE: DISKETTE, 3.50 INCH
52 COMPUTER: IBM PC COMPATIBLE
53 OPERATING SYSTEM: MICROSOFT WINDOWS 95
54 SOFTWARE: MICROSOFT WORD FOR WINDOWS 95 (7.0)
55 CURRENT APPLICATION DATA:
56 APPLICATION NUMBER: US/08/737,524B
57 FILING DATE:
58 CLASSIFICATION: 800
59 ATTORNEY/AGENT INFORMATION:
60 NAME: LYNNE M. CHRISTENBURY
61 REGISTRATION NUMBER: 0,971
62 REFERENCE/DOCKET NUMBER: 1059-A
63 TELECOMMUNICATION INFORMATION:
64 TELEPHONE: 302-952-5491
65 TELEFAX: 302-773-0164
66 TELEX: 835420
67 INFORMATION FOR SEQ ID NO: 27:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 509 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; MOLECULE TYPE: protein
; US-08-737-5248-27

Query Match 16.5%; Score 60.5; DB 2; Length 509;
Best Local Similarity 54.8%; Pred. No. 5;
Matches 17; Conservative 5; Mismatches 8; Indels 1; Gaps 1;

QY 11 GRSGAGSCGTSNDMAAAALASSAAALQII 41
| | | | | : | | | | | | | | | |
Db 71 GRGAR-GVASSHAASAAAAASAAAEVSAI 100

RESULT 7
US-07-792-885A-1
; Sequence 1, Application US/07792885A
; Patent No. 5516651
; GENERAL INFORMATION:
; APPLICANT: Goldring, Steven R.
; APPLICANT: Gorn, Alan H.
; APPLICANT: Lin, Herb Y.
; TITLE OF INVENTION: MAMMALIAN CALCITONIN RECEPTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2864
; COMPUTER: IBM PS/2 Model 502 or 55X
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07792.885A
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; LENGTH: 482
; TYPE: AMINO ACID
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
; US-07-792-885A-1

Query Match 16.5%; Score 59; DB 1; Length 482;
Best Local Similarity 32.4%; Pred. No. 7.4;
Matches 12; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

QY 4 WGEFVQGRSGCGTSDNAAAALASSAAALQI 40
| | | | | : | | | | | | | | | |
Db 407 WNOYQWRWGRSTRANAAATATAAAALAEV 443

; SEQUENCE CHARACTERISTICS:
; LENGTH: 482
; TYPE: AMINO ACID
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
; US-07-792-885A-1
```

```

RESULT 8
US-08-142-439A-7
; Sequence 7, Application US/08142439A
; Patent No. 5670360
; GENERAL INFORMATION:
; APPLICANT: Bernard
; TITLE OF INVENTION: Receptor for the Glucagon-Like-Peptide-1
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5670360 No. 5670360disk of No. 5670360th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/142.439A
; FILING DATE: 24-NOV-93
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 398/92
; FILING DATE: 23-MAR-92
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP93/00697
; FILING DATE: 23-MAR-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Harrington, James J.
; REGISTRATION NUMBER: 38,711
; REFERENCE/DOCKET NUMBER: 3756.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 7:
; LENGTH: 482 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Sus scrofa
; US-08-142-439A-7

Query Match 16.5%; Score 59; DB 1; Length 482;
Best Local Similarity 32.4%; Pred. No. 7.4;
Matches 12; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

QY 4 WGEFVQGRSGCGTSDNAAAALASSAAALQI 40
| | | | | : | | | | | | | | | |
Db 407 WNOYQWRWGRSTRANAAATATAAAALAEV 443

; SEQUENCE CHARACTERISTICS:
; LENGTH: 482
; TYPE: AMINO ACID
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
; US-08-142-439A-7

RESULT 9
US-08-869-477-7
; Sequence 7, Application US/08869477
; Patent No. 5846747
; GENERAL INFORMATION:
; APPLICANT: Bernard
; TITLE OF INVENTION: Receptor for the Glucagon-Like-Peptide-1
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5846747 No. 5846747disk of No. 5846747th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
```





GENERAL INFORMATION: 1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840

```

;
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/881,075
; FILING DATE: 19920511
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: KESNE, PETER H.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 714-154-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 485 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-07-881-075-1

```

```

Query Match          15.28; Score 54.5; DB 1; Length 485;
Best Local Similarity 43.24; Pred. No. 30;
Matches 16; Conservative 6; Mismatches 10; Indels 5; Gaps 1;

Oy 6 EFOVDGRSGRGSCOT-----SNDAAALAAASSAAAA 37
Db 4 DFIMANTGAGGGVDTQALQKMSAAAAAAVAATNAAAA 40

```

```

RESULT 15
US-08-120-827-1
; Sequence 1, Application US/08120827
; Patent No. 5525495
; GENERAL INFORMATION:
; APPLICANT: KESNE, JACK D.
; INVENTOR: KESNE, PETER H.
; APPLICANT: LEVINE, TODD
; TITLE OF INVENTION: METHODS AND COMPOSITIONS USEFUL IN THE
; TITLE OF INVENTION: RECOGNITION, BINDING AND EXPRESSION OF RIBONUCLEIC ACIDS
; TITLE OF INVENTION: INVOLVED IN CELL GROWTH, NEOPLASIA AND IMMUNOREGULATION
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ORION, SPIVAK, MCCLELAND, MAIER & NEUSTADT,
; ADDRESS: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/120,827
; FILING DATE: 15-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KESNE, PETER H.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 714-158-0 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)413-3000
; TELEFAX: (703)413-2220
;
; TELEX: 248855 OPAT UR

```

```

;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 485 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-120-827-1
;
; Query Match          15.28; Score 54.5; DB 1; Length 485;
; Best Local Similarity 43.24; Pred. No. 30;
; Matches 16; Conservative 6; Mismatches 10; Indels 5; Gaps 1;

Oy 6 EFOVDGRSGRGSCOT-----SNDAAALAAASSAAAA 37
Db 4 DFIMANTGAGGGVDTQALQKMSAAAAAAVAATNAAAA 40

```

```

Search completed: February 3, 2001, 02:19:42
Job time: 1704 sec

```

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OK protein - protein search, using sw model

Run on: February 3, 2001, 02:15:26 : Search time 144.12 seconds  
(without alignments)  
16.608 Million cell updates/sec

Title: US-09-289-346a-3

Best score: 359

Sequence: 1 TLVWGFQVQDSRGCCQT.....FQFHNLNSNLDIFDKTPE70

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 segs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq.36.\*

- 1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.\*
- 2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.\*
- 3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.\*
- 4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.\*
- 5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.\*
- 6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.\*
- 7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.\*
- 8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.\*
- 9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.\*
- 10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.\*
- 11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.\*
- 12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.\*
- 13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.\*
- 14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.\*
- 15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.\*
- 16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.\*
- 17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.\*
- 18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.\*
- 19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.\*
- 20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.\*
- 21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Score	Query Match	Length	DB ID	Description
1	218	60.9	361	18 W34336	Tomato mottle virus
2	218	60.9	361	18 W34324	Tomato mottle virus
3	218	60.9	361	18 W34325	Tomato mottle virus
4	218	60.9	361	18 W34326	Tomato mottle virus
5	217	60.6	353	18 W34338	Bean golden mosaic
6	217	60.6	353	18 W34339	Bean golden mosaic
7	217	60.6	353	18 W34332	Bean golden mosaic
8	217	60.6	353	18 W34333	Bean golden mosaic
9	217	60.6	353	18 W34335	Bean golden mosaic
10	215	60.1	353	8 P70407	ORF 4 gene product
11	212	59.2	359	17 R88870	Sardinian tomato y
12	212	59.2	359	17 R88871	Sardinian tomato y

13	212	59.2	359	17 R88872	Sardinian tomato y
14	200.5	56.0	361	8 P70562	Product of ORF 4 f
15	198	55.3	362	18 W56495	Tobacco leaf curl
16	190	53.1	357	18 W34337	Tomato yellow leaf
17	190	53.1	357	18 W34329	Tomato yellow leaf
18	190	53.1	357	18 W34330	Tomato yellow leaf
19	190	53.1	357	18 W34331	Tomato yellow leaf
20	62.5	17.5	579	19 W63707	Human NSK2 protein
21	61.5	17.2	66	11 R08067	Synthetic antifeedant
22	61	17.0	1212	20 W88287	Rat (Na,K)-ATPase
23	61	17.0	1212	20 W88288	Rat (Na,K)-ATPase
24	61	17.0	1212	20 W88289	Rat (Na,K)-ATPase
25	60.5	16.9	131	18 W34327	Tomato yellow leaf
26	60.5	16.9	131	18 W34328	Tomato yellow leaf
27	60.5	16.9	509	17 R85311	Cystathionine gamma
28	60	16.8	314	18 W14283	Human neuroblastoma
29	60	16.8	314	21 P70572	Mouse Phox2b prote
30	59	16.5	3596	21 Y87407	Bordetella pertuss
31	59	16.5	3647	11 R05041	Filamentous haemag
32	58.5	16.3	512	19 W68473	HIV-1 strain YBF30
33	58.5	16.3	512	19 W68474	HIV-1 strain YBF30
34	58.5	16.3	512	19 W68475	HIV-1 strain YBF30
35	57	15.9	293	11 R03623	Zinc finger protein
36	54.5	15.2	46	11 R08098	Synthetic antifeedant
37	54.5	15.2	483	14 R43893	elav. Drosophila
38	54.5	15.2	485	19 W60577	Drosophila neuron
39	54.5	15.2	738	19 W56163	New DNA sequence 1
40	54	15.1	52	20 W69001	Fragment of human
41	54	15.1	52	20 W69002	Fragment of human
42	54	15.1	753	20 W93495	Heat shock factor
43	53.5	14.9	41	11 R08065	Heat shock factor
44	53.5	14.9	73	11 R08066	Synthetic antifeedant
45	53.5	14.9	264	20 Y14820	Synthetic antifeedant

#### ALIGNMENTS

RESULT	1
ID	W34336 Standard; Protein: 361 AA.
XX	W34336;
AC	W34336;
XX	W34336;
DE	27-APR-1998 (first entry)
XX	Tomato mottle virus AC1 protein.
DE	Geminivirus ToMoV; AC1 gene: transdominant mutation;
XX	transgenic plant; disease resistance.
KW	Tomato mottle virus isolate Florida.
XX	W09739110-A1.
XX	W09739110-A1.
XX	W09739110-A1.
PD	23-OCT-1997.
XX	15-APR-1997: 97WO-US06300.
PF	15-APR-1996; 96US-0015517.
XX	16-APR-1996; 96US-0015517.
XX	(SBMI-) SEMINIS VEGETABLE SEEDS INC.
PA	(WISC) WISCONSIN ALUMNI RES FOUND.
XX	Abiquist PG, Hanson SF, Lau HT, Maxwell DP, Stout JT;
XX	WPI: 1997-526447/48.
XX	N-PSDB: T93294.
XX	Transgenic plants expressing geminivirus AC1 and CI wild-type and
PT	mutant genes - have increased resistance to geminivirus infection
PT	e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT	golden mosaic geminivirus

```

XX Example 3.3; Page 57-58; 132pp; English.
XX This protein comprises the wild-type AC1 protein of tomato mottle
XX virus (TOMOV). The AC1 protein is a bifunctional protein. The
XX gene (see T93294) must be expressed for efficient replication of
XX the two genomic components, DNA-A and DNA-B. The AC1 protein has a
XX DNA binding site specific to the DNA-A common region, a DNA nicking
XX activity, and an NTP binding activity. The infection involves
XX production of transgenic plants containing DNA comprising AC1 or CI
XX wild-type or mutant sequences that negatively interfere in trans
XX plants are resistant to viral infection. The AC1/CI transgenic
XX plants are especially resistant to viral infection. The AC1/CI
XX especially from TOMOV, tomato yellow leaf curl virus or bean golden
XX mosaic geminivirus (see T93282-93) and encode polypeptides (see
XX W34324-35) that have mutations in the highly conserved DNA-nicking
XX domain and/or the NTP-binding domain.
XX Sequence 361 AA;

Query Match 60.9%; Score 218; DB 18; Length 361;
Best Local Similarity 57.1%; Pred. No. 28-20;
Matches 40; Conservative 15; Mismatches 15; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDSRGCGQTSNDAAALAAASSAAALQITREKIPKIFOFHNLNSL 60
DB 110 TLVWGEFQVDSRGCGQTSNDAAALAAASSAAALQITREKIPKIFOFHNLNSL 169
QY 61 DRIFOKTPEP 70
DB 170 erifakapep 179

RESULT 2
W34324
ID W34324 standard; Protein; 361 AA.
AC W34324;
XX 27-APR-1998 (first entry)
XX Tomato mottle virus AC1 mutant TOMV-AC1dml.
XX Geminivirus; TOMOV-AC1dml; AC1 gene; transdominant mutation;
XX transgenic plant; disease resistance.
XX Tomato mottle virus isolate Florida.
XX Synthetic.
XX Tomato mottle virus isolate Florida.
XX W09739110-A1.
XX 23-OCT-1997.
XX 15-APR-1997; 97MO-US06300.
XX 16-APR-1996; 96US-0015517.
XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.
XX (WISC-) WISCONSIN ALUMNI RES FOUND.
XX Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX WPI; 1997-526447/48.
XX N-PSDB; T93282.
XX Transgenic plants expressing geminivirus AC1 and CI wild-type and
XX mutant genes - have increased resistance to geminivirus infection
XX e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
XX golden mosaic geminivirus
XX Example 3.4; Page 60-62; 132pp; English.

```

```

CC This protein comprises a transdominant lethal mutant, designated
CC TOMOV-AC1dml, of tomato mottle virus (TOMOV) AC1 protein (see
CC T93294). It is encoded by a mutant AC1 gene (see T93282) of
CC tomato mottle virus (TOMOV), and carries 3 mutations in its
CC expressed domain. The gene (see also T93294) must be
CC expressed efficiently for efficient replication of the
CC DNA-A and DNA-B, of the bipartite TOMOV genome. The infection
CC involves production of transgenic plants containing DNA comprising
CC geminivirus AC1 or CI wild-type or mutant sequences that negatively
CC interfere in trans with geminiviral replication during infection.
CC Such transgenic plants are resistant to viral infection. The
CC AC1/CI genes are especially from TOMOV, tomato yellow leaf curl
CC virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC conserved DNA-nicking and/or NTP-binding domains.
XX Sequence 361 AA;

Query Match 60.9%; Score 218; DB 18; Length 361;
Best Local Similarity 57.1%; Pred. No. 28-20;
Matches 40; Conservative 15; Mismatches 15; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDSRGCGQTSNDAAALAAASSAAALQITREKIPKIFOFHNLNSL 60
DB 110 TLVWGEFQVDSRGCGQTSNDAAALAAASSAAALQITREKIPKIFOFHNLNSL 169
QY 61 DRIFOKTPEP 70
DB 170 erifakapep 179

RESULT 3
W34325
ID W34325 standard; Protein; 361 AA.
AC W34325;
XX 27-APR-1998 (first entry)
XX Tomato mottle virus AC1 mutant TOMV-AC1dml.
XX Geminivirus; TOMOV-AC1dml; AC1 gene; transdominant mutation;
XX transgenic plant; disease resistance.
XX Tomato mottle virus isolate Florida.
XX Synthetic.
XX W09739110-A1.
XX 23-OCT-1997.
XX 15-APR-1997; 97MO-US06300.
XX 16-APR-1996; 96US-0015517.
XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.
XX (WISC-) WISCONSIN ALUMNI RES FOUND.
XX Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX WPI; 1997-526447/48.
XX N-PSDB; T93283.
XX Transgenic plants expressing geminivirus AC1 and CI wild-type and
XX mutant genes - have increased resistance to geminivirus infection
XX e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
XX golden mosaic geminivirus
XX Example 3.5; Page 64-65; 132pp; English.
XX This protein comprises a transdominant lethal mutant, designated
XX TOMOV-AC1dml, of tomato mottle virus (TOMOV) AC1 protein (see

```

NTP-binding domain. The ACL gene (see also Tg3292A), must be expressed for efficient replication of the two genomic components. CCC  
 CCC DNA-A and DNA-B of the bipartite TOMOV genome. The invention CCC  
 CCC involves production of transgenic plants containing DNA comprising CCC  
 CCC geminivirus ACL or C1 wild-type or mutant sequences that negatively CCC  
 CCC interfere in trans with geminiviral replication during infection. CCC  
 CCC Such transgenic plants are resistant to viral infection. CCC  
 CCC ACL/C1 genes are especially from TOMOV, tomato yellow leaf curl CCC  
 CCC virus or bean golden mosaic geminivirus (see Tg3242-93) and encode CCC  
 CCC polypeptides (see W4324-35) that have mutations in the highly CCC  
 CCC conserved DNA-nicking and/or NTP-binding domains. CCC  
 XX Sequence 361 AA:

Query Match 60.9%; Score 218; DB 18; Length 361;  
Best Local Similarity 57.1%; Pred. No. 2e-20;  
Matches 40; Conservative 15; Mismatches 15; Indels  
1 TLWGEFVDRSARGCGCTSSDAAAAAASAAALQITREPKYLPFGHNI  
|||::|||||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
110 tLwGfVdIqRgRGgggsandSyakAlnasvsgslvreepgkdfvglnin  
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
61 DRIKFTKPT 70 :|||::|||  
170 erifakeap 179

RESULT	5
W34338	
ID	W34338 standard; Protein; 353 AA.
XX	
ID	W34338;
XX	
XX	27-APR-1998 (first entry)
XX	
DE	Bean golden mosaic geminivirus CI protein.
XX	

OS Bean golden mosaic virus type II isolate Guatemala.

PN W09739110-A1.

23-OCT-1997.

PF 15-APR-1997; 97WO-US06300.

PR 16-APR-1996; 96US-0015517.

PA (SEMI-) SEMINIS VEGETABLE SEEDS INC.

XX  
XX

PI Ahlquist PG, Hanson

DR N-PSDB; T93314.

PT Transgenic plants expressing geminivirus AC1 and CI wild-type and mutant genes - have increased resistance to geminivirus infection e.g. tomato mottle virus, tomato yellow leaf curl virus or bean

PT golden mosaic geminivirus

XX

PS Example 5; Page 100-102; 132pp; English.  
XX  
CCC This sequence comprises the bean golden mosaic virus (BGMV) C1

protein that is required for replication. The invention involves production of transgenic plants containing DNA comprising CI or AC1 wild-type or mutant sequences that negatively interfere in trans

CC with geminiviral replication during infection. Such transgenic

CC plants are resistant to viral infection. The AC1/C1 genes are

CC curl virus (see T93282-93) and encode polypeptides (see W34324-35)  
 CC that have mutations in the highly conserved DNA-nicking domain  
 CC and/or the NTP-binding domains.

XX Sequence 353 AA;

Query Match 60.6%; Score 217; DB 18; Length 353;

Best Local Similarity 58.6%; Pred. No. 2.6e-20;

Matches 41; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

OY 1 TLVNGFEVDGSRGCGCTSDNDAAAALAAASSAAALQIIREKIPKXLYLQFINLSNL 60

Db 110 TLVNGFEVDGSRGCGCTSDNDAAAALAAASSAAALQIIREKIPKXLYLQFINLSNL 169

OY 61 DRIFOKTPEP 70

Db 170 erifvkyppep 179

RESULT 6

W34332  
 ID W34332 standard; Protein; 353 AA.

XX W34332;

XX 27-APR-1998 (first entry)

XX Bean golden mosaic geminivirus C1 BGA190 mutant.

XX Geminivirus; BGWV; C1 gene; transdominant mutation;

XX transgenic plant; disease resistance.

XX Bean golden mosaic virus type II isolate Guatemala.

XX W09739110-AL.

XX 23-OCT-1997.

XX 15-APR-1997; 97WO-US06300.

XX 16-APR-1996; 96US-0015517.

XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.

XX (WISC ) WISCONSIN ALUMNI RES FOUND.

XX Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;

XX WPI; 1997-526447/48.

XX N-PSDB; T93290.

XX Transgenic plants expressing geminivirus AC1 and C1 wild-type and  
 XX mutant genes - have increased resistance to geminivirus infection  
 XX e.g. tomato mottle virus, tomato yellow leaf curl virus or bean  
 XX golden mosaic geminivirus

XX Example 5: Page 103-105; 132pp; English.

XX This protein comprises a control mutant of the bean golden mosaic  
 XX virus (BGWV) C1 protein (see W34338) that is required for  
 XX replication. It is encoded by mutated C1 open reading frame

XX BGAC190 (see T93290). The invention involves production of

XX transgenic plants containing DNA comprising geminivirus C1 or AC1

XX wild-type or mutant sequences that negatively interfere in trans

XX with geminiviral replication during infection. Such transgenic

XX plants are resistant to infection by the AC1 genes and/or C1

XX genes especially from BGWV, tomato mottle virus or tomato leaf

XX curl virus (see T93282-93) and encode polypeptides (see W34324-35)

XX that have mutations in the highly conserved DNA-nicking and/or the

XX NTP-binding domains.

XX Sequence 353 AA;

Query Match 60.6%; Score 217; DB 18; Length 353;

Best Local Similarity 58.6%; Pred. No. 2.6e-20;

Matches 41; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

OY 1 TLVNGFEVDGSRGCGCTSDNDAAAALAAASSAAALQIIREKIPKXLYLQFINLSNL 60

Query Match 60.6%; Score 217; DB 18; Length 353;

Best Local Similarity 58.6%; Pred. No. 2.6e-20;

Matches 41; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

OY 1 TLVNGFEVDGSRGCGCTSDNDAAAALAAASSAAALQIIREKIPKXLYLQFINLSNL 60

Db 110 TLVNGFEVDGSRGCGCTSDNDAAAALAAASSAAALQIIREKIPKXLYLQFINLSNL 169

OY 61 DRIFOKTPEP 70

Db 170 erifvkyppep 179

RESULT 7

W34333  
 ID W34333 standard; Protein; 353 AA.

XX W34333;

XX 27-APR-1998 (first entry)

XX Bean golden mosaic geminivirus C1 BGA221 mutant.

XX Geminivirus; BGWV; C1 gene; transdominant mutation;

XX transgenic plant; disease resistance.

XX Bean golden mosaic virus type II isolate Guatemala.

XX W09739110-AL.

XX 23-OCT-1997.

XX 15-APR-1997; 97WO-US06300.

XX 16-APR-1996; 96US-0015517.

XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.

XX (WISC ) WISCONSIN ALUMNI RES FOUND.

XX Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;

XX WPI; 1997-526447/48.

XX N-PSDB; T93291.

XX Transgenic plants expressing geminivirus AC1 and C1 wild-type and  
 XX mutant genes - have increased resistance to geminivirus infection  
 XX e.g. tomato mottle virus, tomato yellow leaf curl virus or bean  
 XX golden mosaic geminivirus

XX Example 5: Page 107-109; 132pp; English.

XX This protein comprises a transdominant lethal mutant of the bean  
 XX golden mosaic virus (BGWV) C1 protein (see W34338) that is required  
 XX for replication. It is encoded by mutated C1 open reading frame

XX BGAC221 (see T93291) and carries a mutation in the NTP-binding

XX domain. The invention involves production of transgenic plants

XX containing DNA comprising geminivirus C1 or AC1 wild-type or mutant

XX genes. Such transgenic plants are resistant to infection by the

XX AC1 genes and/or C1 genes especially from BGWV, tomato mottle

XX virus or tomato leaf curl virus (see T93282-93)

XX and encode polypeptides (see W34324-35) that have mutations in the

XX highly conserved DNA-nicking and/or the NTP-binding domains.

XX Sequence 353 AA;



RESULT	9	
ID	W34335	
CD	W34335 standard; Protein; 353 AA.	
CC	W34335;	
CC	W34335;	
DT	27-APR-1998 (first entry)	
DE	Bean golden mosaic geminivirus Cl BGA262 mutant.	
KW	Geminivirus; BGWV; Cl gene; transdominant mutation;	
KW	transgenic plant; disease resistance.	
OS	Bean golden mosaic virus type II isolate Guatemala.	
PN	W09739110-A1.	
XX		
XX	23-OCT-1997.	
PF		
PF	15-APR-1997: 97WU-USO6300.	
PR		
PR	16-APR-1996: 96US-0015517.	
XX		
XX	(SEMI-) SEMINIS VEGETABLE SEEDS INC.	
PA	(WISC-) WISCONSIN ALUMNI RES FOUND.	
XX		
XX	Ahlquist PG, Hanson SP, Luu HT, Maxwell DP, Stout JT;	
XX	WPI: T97-526447/48.	
DR	N-PSDB: T93293.	
XX		
XX	Transgenic plants expressing geminivirus ACl and Cl wild-type and	
PT	mutant genes - have increased resistance to geminivirus infection	
PT	e.g. tomato mottle virus, tomato yellow leaf curl virus or bean	
PT	golden mosaic geminivirus	
XX		
XX	Example 5; Page 115-116; 132pp; English.	
CC		
CC	This protein comprises a transdominant lethal mutant of the bean	
CC	golden mosaic virus (BGWV) Cl protein (see W34338) that is required	
CC	for replication; it is encoded by mutated Cl open reading frame	
CC	2392-2402 nt. The protein is a transdominant lethal mutant of the	
CC	domain. The invention involves production of transgenic plants	
CC	containing DNA comprising geminivirus Cl or ACl wild-type or mutant	
CC	sequences that negatively interfere in trans with geminiviral	
CC	replication during infection. Such transgenic plants are resistant	
CC	to viral infection. The ACl/Cl genes are especially from BGWV.	
CC	CC tomato mottle virus or tomato yellow leaf curl virus (see T93282-93).	
CC	CC transgenic polyploid plants (see Example 2) mutants in the	
CC	CC highly conserved DNA-nicking and/or the NTP-binding domains.	
CC		
SQ	Sequence 353 AA;	
Query Match	60.6%; Score 217; DB 18; Length 353;	
Residual Similarity	18.8%;	
Matches	41; Conservative 16; Mismatches 16; Indels 0; Gaps	

	RESULT	10
Db	110	tfwqgfydvgrsggqaaandsyaka.hnadsa.tllkseeqkyvqlhmlrsl
QY	61	DRIFDKTPPE 70
Db	170	erfivkvpep 179
	RESULT	10
XX	170	70407
XX	170	P70407 standard; Protein; 353 AA.

```

AC P70407;
XX 02-MAY-1991 (first entry)
XX
XX DE ORF 4 gene product of Bean Golden Mosaic virus.
XX
XX KW Plant vector.
XX
XX OS Bean golden mosaic virus.
XX
XX XX JP61257186-A.
XX
XX PD 14-NOV-1986.
XX
XX PF 10-MAY-1985; 85JP-0098108.
XX
XX PR 10-MAY-1985; 85JP-0098108.
XX
XX XX (TEIJ ) TEIJIN KK.
XX
XX WPI: 1987-159662/23.
XX
XX DR N-PSDB: N70630.
XX
XX PT New DNA and hybrid DNA - used for recombinant vector of plants.
XX
XX PS Disclosure: Fig 6; 24pp; Japanese.
XX
XX The sequence encoding this protein may be taken along with the -a
XX DNA sequence and a suitable resistance gene, and used to create a
XX recombinant plant vector.
XX
XX See also N70629.
XX
XX Sequence 353 AA;

Query Match 60.1%; Score 215; DB 8; Length 353;
Best Local Similarity 58.6%; Pred. No. 4.8e-20;
Matches 41; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLVNGEFQVDSARGGCGTSDNAAAALAAASAAALQITREKIPKYLQFHNLSNL 60
Db 110 TLWGFQVDSARGGCGTSDNAAAALAAASAAALQITREKIPKYLQFHNLSNL 169
QY 61 DRIFDKTPEP 70
Db 170 eriffkvpep 179

RESULT 11
ID R88871 standard; Protein: 359 AA.
XX
XX AC R88870;
XX
XX DT 07-NOV-1996 (first entry)
XX
XX DE Sardinian tomato yellow leaf curl virus mutated Rep protein (K227A).
XX
XX KW Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
XX modification; mutation; viral replication; deficient; inhibition;
XX viral resistance; geminivirus; tomato yellow leaf curl virus;
XX Sardinian isolate; SYICV; transgenic plant; P-loop; C1 protein;
XX A11 protein; dominant negative phenotype.
XX
XX OS Sardinian tomato yellow leaf curl virus.
XX
XX FT Key Location/Qualifiers
XX PH Misc-difference 227
XX /note= "wild-type Lys has been replaced by Ala"
XX
XX PN W09608573-AL.
XX
XX PD 21-MAR-1996.
XX
XX PF 15-SEP-1995; 95WO-FR01192.
XX
XX PR 15-SEP-1994; 94FR-0011040.

```

```

XX 15-SEP-1995; 95WO-FR01192.
XX
XX 15-SEP-1994; 94FR-0011040.
XX
XX (CNRS ) CENT NAT RECH SCI.
XX
XX Gronenborn B;
XX
XX WPI: 1996-179947/18.
XX
XX DR N-PSDB: T12904.
XX
XX Prodn. of virus-resistant transgenic plants - using mutated genomic
XX PT sequence from phytopathogenic DNA virus
XX
XX PS Disclosure: Fig 13; 93pp; French.
XX
XX CC Mutation of consensus amino acids in the NTP-binding site of
XX geminivirus Rep protein is used to produce replication deficient
XX viruses; the mutated viral nucleic acid is used for producing
XX transgenic plants that are resistant to geminiviruses; alternative
XX CC virus; the present sequence is a mutant form of the Rep (or C1)
XX protein from the Sardinian isolate of tomato yellow leaf curl virus
XX (SYICV) in which the wild-type Lys227 residue has been changed to an
XX CC Ala residue; transgenic Nicotiana benthamiana plants generated by
XX transformation with the mutated virus were found to be resistant to
XX SYICV, i.e. the mutation results in a dominant negative phenotype.
XX
XX Sequence 359 AA;

Query Match 59.2%; Score 212; DB 17; Length 359;
Best Local Similarity 56.5%; Pred. No. 1.2e-19;
Matches 39; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

QY 2 LVNGEFQVDSARGGCGTSDNAAAALAAASAAALQITREKIPKYLQFHNLSNL 61
Db 111 LWGTFQVDSARGGCGTSDNAAAALAAASAAALQITREKIPKYLQFHNLSNL 170
QY 62 RIFDKTPEP 70
Db 171 kvfqppep 179

RESULT 12
ID R88871 standard; Protein: 359 AA.
XX
XX AC R88871;
XX
XX XX 07-NOV-1996 (first entry)
XX
XX DE Sardinian tomato yellow leaf curl virus mutated Rep protein (K227H).
XX
XX KW Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
XX modification; mutation; viral replication; deficient; inhibition;
XX KW viral resistance; geminivirus; tomato yellow leaf curl virus;
XX Sardinian isolate; SYICV; transgenic plant; P-loop; C1 protein;
XX A11 protein.
XX
XX OS Sardinian tomato yellow leaf curl virus.
XX
XX FT Key Location/Qualifiers
XX PH Misc-difference 227
XX /note= "wild-type Lys has been replaced by His"
XX
XX PN W09608573-AL.
XX
XX PD 21-MAR-1996.
XX
XX PF 15-SEP-1995; 95WO-FR01192.
XX
XX PR 15-SEP-1994; 94FR-0011040.

```

```

XX (CNRS ) CENT NAT RECH SCI.
PA Gronenborn B;
XX WPI: 1996-179947/18.
XX N-PSDB: T12905.
DR Prodn. of virus-resistant transgenic plants - using mutated genomic
PT sequence from phytopathogenic DNA virus
XX Disclosure: Fig 13; 93pp; French.
XX Mutation of consensus amino acids in the NTP-binding site of
CC geminivirus Rep protein is used to produce replication deficient
CC viruses. The mutated viral nucleic acid is used for producing
CC transgenic plants that are resistant to, or tolerant of, the native
CC virus. The present sequence is a mutant form of the Rep (or C1)
CC protein from the Sardinian isolate of tomato yellow leaf curl virus
CC (STYLVCV) in which the wild-type Lys227 residue has been changed to
CC an Arg residue; transgenic Nicotiana benthamiana plants generated by
CC transformation with the mutated virus were not resistant to STYLVCV.
CC In contrast, plants transformed with a virus in which Lys227 had been
CC replaced by Ala were found to be resistant.
XX Sequence 359 AA:

Query Match 59.2%; Score 212; DB 17; Length 359;
Best Local Similarity 56.5%; Pred. No. 1.2e-19;
Matches 39; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

QY 2 LVNGEFQVDRSGAGCQTSNDAAAALAAASSAAALQIIREKIPEKYLFOFHNLNSLND 61
DB 111 Lewtfgldrgsarggggtandayakainagskaldvikelaprdyvlhfhnsnld 170
QY 62 RIFDKTPPEP 70
DB 171 kvfqvppap 179

RESULT 13
ID R88872 standard; Protein: 359 AA.
XX R88872;
XX R88872;
XX 07-NOV-1996 (first entry)
XX Sardinian tomato yellow leaf curl virus mutated Rep protein (#227R).
XX Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
XX modification; mutation; viral replication; deficient; inhibition;
XX viral resistance; geminivirus; tomato yellow leaf curl virus;
XX Sardinian isolate; STYLVCV; transgenic plant; P-loop; C1 protein;
XX AL1 protein.
XX Sardinian tomato yellow leaf curl virus.
XX key Location/Qualifiers
XX Misc-difference 227
XX /note= "wild-type Lys has been replaced by Arg"
XX WO9608573-A1.
XX 21-MAR-1996.
XX 15-SEP-1995; 95WO-FR01192.
XX 15-SEP-1994; 94FR-0011040.
XX (CNRS ) CENT NAT RECH SCI.

```

```

PI Gronenborn B;
XX WPI: 1996-179947/18.
XX N-PSDB: T12906.
DR Prodn. of virus-resistant transgenic plants - using mutated genomic
PT sequence from phytopathogenic DNA virus
XX Disclosure: Fig 13; 93pp; French.
XX Mutation of consensus amino acids in the NTP-binding site of
CC geminivirus Rep protein is used to produce replication deficient
CC viruses. The mutated viral nucleic acid is used for producing
CC transgenic plants that are resistant to, or tolerant of, the native
CC virus. The present sequence is a mutant form of the Rep (or C1)
CC protein from the Sardinian isolate of tomato yellow leaf curl virus
CC (STYLVCV) in which the wild-type Lys227 residue has been changed to
CC an Arg residue; transgenic Nicotiana benthamiana plants generated by
CC transformation with the mutated virus were not resistant to STYLVCV.
CC In contrast, plants transformed with a virus in which Lys227 had been
CC replaced by Ala were found to be resistant.
XX Sequence 359 AA:

Query Match 59.2%; Score 212; DB 17; Length 359;
Best Local Similarity 56.5%; Pred. No. 1.2e-19;
Matches 39; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

QY 2 LVNGEFQVDRSGAGCQTSNDAAAALAAASSAAALQIIREKIPEKYLFOFHNLNSLND 61
DB 111 Lewtfgldrgsarggggtandayakainagskaldvikelaprdyvlhfhnsnld 170
QY 62 RIFDKTPPEP 70
DB 171 kvfqvppap 179

RESULT 14
ID P70562 standard; Protein: 361 AA.
XX P70562;
XX P70562;
XX 30-APR-1991 (first entry)
XX Product of ORF 4 from MYMV complementary strand (b).
XX Geminivirus.
XX Mungbean yellow mosaic virus.
XX Jp62126982-A.
XX 09-JUN-1987.
XX 28-NOV-1985; 85JP-0266080.
XX 28-NOV-1985; 85JP-0266080.
XX (TEIJ ) TEIJIN KK.
XX WPI: 1987-196308/28.
XX N-PSDB: N70895.
XX Novel DNA and hybrid DNA useful - as vector for recombinant work
XX of plant gene.
XX Disclosure: Fig 8; 21pp; Japanese.
XX The sequence is encoded by ORF 4 which occurs on the complementary
XX strand of the (b) molecule of the geminivirus.
XX See also P70559-P70567.

```

Search completed: February 3, 2001, 02:15:27  
Job time: 3434 sec

GenCode version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.  
OW protein - protein search, using sw model  
Run on: February 3, 2001, 02:22:50 ; Search time 160.33 Seconds  
(without alignments)  
45.497 Million cell updates/sec

Title: US-09-289-3464-4  
Perfect score: 361  
Sequence: 1 TLVWGFQVDSRGSGCQT.....QFHNLSNLDRIEDXTEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 segs, 117027915 residues  
Total number of hits satisfying chosen parameters: 374700  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL15:.\*  
1: sp.archaea:.\*  
2: sp.bacteria:.\*  
3: sp.fungi:.\*  
4: sp.human:.\*  
5: sp.invertebrate:.\*  
6: sp.mammalia:.\*  
7: sp.mca:.\*  
8: sp.organelle:.\*  
9: sp.phase:.\*  
10: sp.plant:.\*  
11: sp.podent:.\*  
12: sp.virus:.\*  
13: sp.vertibrate:.\*  
14: sp.unclassified:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	281	77.8	226	12 009727	O09727 leonurus mo
2	281	77.8	226	12 09WHF6	O9WHF6 tomato mild
3	279	77.3	361	12 06W574	O6W574 bean golden
4	277	76.7	225	12 05QDB1	O5QDB1 cowpea gold
5	277	76.7	185	12 09B83	O9B83 sweet potato
6	271	75.1	195	12 09B83	O9B83 sweet potato
7	266	73.7	149	12 09B975	O9B975 macrophyllum
8	266	73.7	233	12 09YL44	O9YL44 macrophyllum
9	260	72.0	234	12 039180	O39180 potato yell
10	254	70.4	190	12 09Z089	O9Z089 tobacco lea
11	254	70.4	190	12 09Z089	O9Z089 tobacco lea
12	253	70.1	180	12 09W827	O9W827 tobacco lea
13	253	69.5	208	12 09Z0C4	O9Z0C4 tobacco lea
14	253	69.5	208	12 09Z0C4	O9Z0C4 tobacco lea
15	251	69.5	208	12 09Z088	O9Z088 tobacco lea
16	251	69.5	363	12 072705	O72705 cotton leaf
17	251	69.5	363	12 072719	O72719 cotton leaf
18	249	69.0	208	12 09Z0B6	O9Z0B6 tobacco lea
19	248	68.7	203	12 09Z083	O9Z083 tobacco lea

# SUMMARIES

09YZV4 tomato yell  
09YZV2 tomato yell  
09YUX7 tomato yell  
09YU77 tomato yell  
09YU77 tomato yell  
089142 tomato yell  
091et7 cotton leaf  
072723 cotton leaf  
09in48 okra enatio  
088888 tomato pscu  
056816 chayote mos  
09Z0A0 tobacco lea  
09Z0A0 tobacco lea  
09Z0B6 tobacco lea  
09Z0C6 tobacco lea  
091eaz cassava dem  
073494 okra yellow  
091et1 cotton leaf  
091et1 beet curly  
072632 beet curly  
072632 beet curly  
09Z0A7 tobacco lea  
09Z0A7 tobacco lea  
072707 cotton leaf  
072710 cotton leaf  
065317 ageratum ye  
09YU77 althea rose

## ALIGNMENTS

RESULT 1  
009727 PRELIMINARY; PRF: 226 AA.  
ID 009727  
AC O09727 1997 (TREMEL:el. 04, Created)  
DT 01-JUL-1997 (TREMEL:el. 04, Last sequence update)  
DT 01-OCT-2000 (TREMEL:el. 15, Last annotation update)  
DE REP PROTEIN (FRAGMENT).  
GN REP.  
OS Leonurus mosaic virus.  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID:58177;  
RN 1  
RC STRAIN:LEMV- BRAZIL 1;  
RL Faria J.C., Maxwell D.P.:  
Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; U92532; AAB51157.1; -;  
DR INTERPRO; IPR001191; -;  
DR INTERPRO; IPR001301; -; 1;  
DR TRAFK; PF00055; Geminiviridae; -;  
DR PRINTS; PR00227; GEMCOATV1.1;  
DR PRINTS; PR00228; GEMCOATV1.1;  
FT NON\_TER 226  
SQ SEQUENCE 226 AA; 25617 MW; 73CDB6E76603FC5 CRC64;

Query Match 77.8%; Score 281; DB 12; Length 226;  
Best Local Similarity 76.6%; Pos. 2, 9e-25;  
Matches 55; Conservative 3; Mismatches 12; Indels 0;  
Gaps 0;

OY 1 TLVWGFQVDSRGSGCQTSDAAAEALNASSKEAQIITAAIPKTYLQFHNLSNLD 60  
DB 111 TVWGFQVDSRGSGCQTSDAAAEALNAPDKRTALQITKEKLPKTYLQFHNLSNLD 170  
OY 61 DRIEDXTEP 70  
DB 171 DRIEDXTEP 180

RESULT 2  
09WHF6

```

ID Q5WHF6 PRELIMINARY: PRT: 226 AA.
AC Q5WHF6; 1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN REP.
OS tomato mild mottle geminivirus.
OC Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.
OX NCBI_Taxid=92943;
RN [1]
RC STRAIN=HNSG-HSKG.
RA Nakhla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,
RT Maxwell D.P.;
RT "Molecular characterization and DNA-based detection methods for
RT vegetable-infecting geminiviruses in Central America.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF131071; A033471.1; -.
DR INTERPRO: IPR001191; -.
DR PFAM: PF00799; Gemin_Ali.1.
DR PRINTS: PR00227; GEMCOATALL.
DR PRINTS: PR00228; GEMCOATCLVLL.
FT NON_TER 226 226
SQ SEQUENCE 226 AA: 25941 MW; 2E4116712871A23 CRC64;

Query Match 77.6%; Score 281; DB 12; Length 226;
Best Local Similarity 74.3%; Pred. No. 2, 9e-25;
Matches 52; Conservative 11; Mismatches 7; Indels 0; Gaps 0;

QY 1 TLVGFQVDSRGSGCOTSDAAAEALNASSKEEALQITAAAIPEKYLFOFHNLSNL 60
Db 111 TLVGFQVDSRGSGCOTSDAAAEALNASSKEEALQITAAAIPEKYLFOFHNLSNL 170
QY 61 DRIFDKTPEP 70
Db 171 DRIFAKAPEP 180

RESULT 3
ID Q67574 PRELIMINARY: PRT: 361 AA.
AC Q67574;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)
DE POTATIVE REPLICATIVE PROTEIN.
GN Ali.
OS golden mosaic virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_Taxid=10839;
RN [1]
RC SEQUENCE FROM N.A.
RA Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,
RA Maxwell D.P., Russell D.R.;
RL Phytopathology 81:980-985(1991).
RC SEQUENCE FROM N.A.
RA Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,
RA Morales F.J., Maxwell D.P.;
RL Plant Dis. 75:336-342(1991).
RN [3]
RC SEQUENCE FROM N.A.
RA Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: M88666; A446312.1; -.
DR INTERPRO: IPR001191; -.
DR PFAM: PF00799; Gemin_Ali.1.
DR PRINTS: PR00227; GEMCOATALL.
DR PRINTS: PR00228; GEMCOATCLVLL.

```

```

DR PRODM: P000736; -. 1.
SQ SEQUENCE 361 AA: 41041 MW; 0094C7ACAF06B788 CRC64;

Query Match 77.3%; Score 279; DB 12; Length 361;
Best Local Similarity 77.6%; Pred. No. 8, 5e-25;
Matches 52; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 4 MGEFQVDSRGSGCOTSDAAAEALNASSKEEALQITAAAIPEKYLFOFHNLSNL 63
Db 113 MGEFQVDSRGSGCOTSDAAAEALNASSKEEALQITAAAIPEKYLFOFHNLSNL 172
QY 64 FKXTPPEP 70
Db 173 FKXAPPEP 179

RESULT 4
ID Q9QDRI PRELIMINARY: PRT: 225 AA.
AC Q9QDRI;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)
DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN REP.
OS cowpea golden mosaic geminivirus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_Taxid=49263;
RN [1]
RC STRAIN=CGMV-BR.
RA Faria J.C.;
RT "Partial nucleotide sequence of cowpea golden mosaic geminivirus from
RT Brazil."; (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF148708; A406318.1; -.
DR INTERPRO: IPR001191; -.
DR PFAM: PF00799; Gemin_Ali.1.
DR PRINTS: PR00227; GEMCOATALL.
DR PRINTS: PR00228; GEMCOATCLVLL.
DT NON_TER 225 225
SQ SEQUENCE 225 AA: 25766 MW; 1089CB5B08015B50 CRC64;

Query Match 76.7%; Score 277; DB 12; Length 225;
Best Local Similarity 77.6%; Pred. No. 8, 4e-25;
Matches 52; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 4 MGEFQVDSRGSGCOTSDAAAEALNASSKEEALQITAAAIPEKYLFOFHNLSNL 63
Db 113 MGEFQVDSRGSGCOTSDAAAEALNASSKEEALQITAAAIPEKYLFOFHNLSNL 172
QY 64 FKXTPPEP 70
Db 173 FKXPEP 179

RESULT 5
ID Q9Q855 PRELIMINARY: PRT: 364 AA.
AC Q9Q855;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
GN AC1
OS sweet potato leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_Taxid=100755;
RN [1]
RC SEQUENCE FROM N.A.

```



	Matches	49;	Conservative	11;	Mismatches	10;	Indels	0;	Gaps	0;
QY	1	TLWCEGFQVDRSGARGCOT	SNDAAAEALNASKEEA	LIITAAAI	PKYKLYQFHNLSNL	60				
DB	110	TIEMWQVQIDRSGARGCOT	SNDAAAEALNASKEEA	LIITAAAI	PKYKLYQFHNLSNL	169				
QY	61	DRIDFNTPEP	70							
DB	170	DRIDFNMKEP	179							
RESULT	9									
Q39180										
ID	Q39180		PRELIMINARY;		PRG:	234	AA.			
AC	Q39180									
DT	01-JAN-1998		(TrEMBLrel. 05, Created)							
DT	01-JAN-1998		(TrEMBLrel. 05, Last sequence update)							
DT	01-OCT-2000		(TrEMBLrel. 15, Last annotation update)							
DT	02		REPLICATION ASSOCIATED PROTEIN (FRAGMENT).							
OC	Q39180		Viruses; ssDNA viruses; Geminiviridae; Begomovirus.							
OC	Q39180		NCBI_TaxID=10827;							
RP	11		SEQUENCE FROM N.A.							
RP	11		STRAIN-TOMATO STRAIN;							
RC	Guzman P., Arredondo C.R., Emmatty D., Portillo R.J., Gilbertson R.L.;									
RL	Plant Dis. 81:312-312(1997).									
RL	EMBL; AF026553; AAB92605.1;									
DR	INTERPRO; IPRO01191;									
DR	IPRO01191; 1;									
DR	PFAM; PF00759; GEMCALF1.1;									
DR	PRINTS; PR00227; GEMCOATL1;									
DR	PRINTS; PR00228; GEMCOATL1L1.									
DR	NON_TER	234								
EQ	SEQUENCE	234	AA;	26486	MW;	9ED8F0697105CD19	CR654;			

Query Match	72.08	Score 260	DB 12	Length 234
Best Local Similarity	68.68	Pos Neg	6.7e-23	
Matches 48	Conservative 12	Mismatches 10	Indels 0	Gaps 0

```

RESULT 10
092089
ID 092089; PRELIMINARY; PRT: 190 AA.
AC 092089;
01-MAY-1999 (TREMBLrel.10. Created)
DT 01-MAY-1999 (TREMBLrel.10. Last sequence update)
DT 01-MAY-2000 (TREMBLrel.13. Last annotation update)
DE 01-MAY-2000 (TREMBLrel.13. Last annotation update)
DE 01-MAY-2000 (TREMBLrel.13. Last annotation update)
DE PARTIAL AND COMPLETE CDS (FRAGMENT).
DE C1.
GN tobacco leaf curl virus.
OC viruses: ssRNA viruses: Geminiviridae: Begomovirus.
NCBI_TaxID=57762;
NCBI_Accession=U00001;
RE SEQUENCE FROM N.A.
RC STRAIN=KOKOHAMA3.
RA Ooi K., Ohshita S., Ishii I., Yahara T.:
RT "Molecular phylogeny of geminivirus infecting wild plants in Japan."
RL J. Plant Res. 110:247-257(1997).
RL EMBL: AB001150; F0434033.1;
DR EMBL: AB001150; F0434033.1;
DR INTERPRO: IPRO01301;
DR PFAM: PF00799; Gemin1-ALL; 1.
DR

```

	DR	PRINTS:	PR00227; GEMCONTALI.	-
	DR	PRINTS:	PR00228; GEMCOTALVL.	-
	FT	NON_TER	1	-
	FT	NON_FT	190	-
	SQ	SEQUENCE	190 AA; 21432 MW; AAC093DID161OFAD CRC64;	-
		Query Match	70.4%; Score 254; DB 12; Length 190;	
		Best Local Similarity	59.8%; Pref No. 3, 5-22;	
		Matches	50; Conservative 8; Mismatches 11; Indels 16; Gaps 1;	
Oy	1	TLVNGEPVQDGSRGCOTSDNAAEALINNAKEEAQLITAAATPKYLFQHILNSU	60	
Dy	85	TLEWTFQIDGRSGGCGNQANDKAEALINNAKAELAIREKLPIKFQHYHLNSU	144	
Oy	61	-----FKPTSE 69		
Dy	145	DRIFAPPLVPVCPTFASTDOVEP 169		

RESULT	11
Q3Z084	
ID	Q3Z084 PRELIMINARY: PRT: 190 AA.
AC	Q3Z084;
DT	01-MAY-1999 (TREMBLrel. 10. Created)
DD	01-MAY-1999 (TREMBLrel. 13. Last sequence update)
DI	01-MAY-2000 (TREMBLrel. 13. Last annotation update)
CL	C1 AND C4 GENES: GLOBE YOKOHAM5-2.
DE	PARTIAL AND COMPLETE CDS (FRAGMENT).
GC	.C1.
GN	Tobacco leaf curl virus.
OS	Viruses; ssDNA viruses: Geminiviridae; Begomovirus.
OC	Viruses; ssDNA viruses: Geminiviridae; Begomovirus.
RL	tbl_xxxid=57762;
RM	11.
RC	SEQUENCE FROM N.A.
RP	STRAIN-YOKOHAMA5;
RR	Ooi K., Ohshita S., Ishii I., Yahara T.;
RA	"Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL	J. Plant Res. 110:247-257(1997).
RD	EMBL accession number: AF034059.1;
DR	INTERPRO: IPRO01301;
DR	INTERPRO: IPRO01301;
DR	PFAM: PF00799; Gmim1c_11; 1.
DR	PRINTS: PR00227; GEMCOATL1.
DR	PRINTS: PR00228; GEMCOATLVL.
DR	NON-TER
FT	190 190
FE	SEQUENCE 190 AA: 21444 MW: 84612GALIFKQLAD_CSC4C
SO	

Query Match	70.4%	Score 254:	DB 12:	length 190:
Similarity	70.4%	Free 25:	35e-42	
Matches	50:	Conservative	8:	Mismatches 11: Indels 16: Gaps 1:
QY 1 TLWCERQVDRGARGCGCTSDNAAEALNASSKEEALQITAAATPKYLYFPHNLNSNI 60				
DB 85 TLEWGTQIDRGARGCGQNDAKAEALNASSKAEALATIRKUPKDTFYQHNLSNI 144				
QY 61 DRL-----FQKTFE 69				
DB 145 DRIEAPLEVPFCVFATSSFDQVPE 169				
RESULT 12				
QWZ 094827				
AC 094827				
DC 01-NOV-1999 (TrEMBLrel. 12, Created)				
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)				
DF 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)				
DE C1 PROTEIN (FRAGMENT)				
OS TOBACCO leaf curl virus.				



OC Viruses: ssDNA viruses: Geminiviridae; Begomovirus.

```

RN (1) SEQUENCE FROM N.A.
RC STRAIN-CORAL1
RA Ooi K., Ohshita S., Ishii I., Yahara T.;
RT "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL J. Plant Res. 110:247-257(1997).
DR EMBL: AB001303; BAA34010.1; -.
DR INTERPRO: IPRO01191; -.
DR PFAM: PF00799; Gemin1_AL1; 1.
DR PRINTS: PR00228; GEMCOATLVL.
DR NON_TER 1 190
FT NON_TER 190 190
SQ SEQUENCE 190 AA; 21444 MW; 93C3742A8EBDB7EB CRC64;

```

```

Query Match 70.1%; Score 253; DB 12; Length 190;
Best Local Similarity 58.8%; Pred. No. 4.5e-22;
Matches 50; Conservative 8; Mismatches 11; Indels 16; Gaps 1;
OY 1 TLVNGEFQVDCRSAGCGCOTSDNAAREALNASSKEEALQIIAAATPEKYLQFQHNLSNL 60
DB 85 TLWNGEFQVDCRSAGCGCQNDACAEALNASSKEEALQIIAAATPEKYLQFQHNLSNL 144
OY 61 DRI-----FCKTPE 69
DB 145 DRIAPLELVFVCPSSSSFDQVPE 169

```

```

RESULT 13
OQ20C4 ID OQ20C4 PRELIMINARY; PRT; 208 AA.
AC OQ20C4
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE C1 AND C4 GENES, CLONE ABUR3-1, PARTIAL
DE PARTIAL AND COMPLETE CDS (FRAGMENT).
GN Tobacco leaf curl virus.
OS Viruses: ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=67762;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-ABUR3
RA Ooi K., Ohshita S., Ishii I., Yahara T.;
RT "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL J. Plant Res. 110:247-257(1997).
DR EMBL: AB001294; BAA33992.1; -.
DR INTERPRO: IPRO01191; -.
DR PFAM: PF00799; Gemin1_AL1; 1.
DR PRINTS: PR00227; GEMCOATLVL.
DR NON_TER 1 190
FT NON_TER 190 190
SQ SEQUENCE 208 AA; 23526 MW; 249CC31D8729C72D CRC64;

```

```

Query Match 70.1%; Score 253; DB 12; Length 208;
Best Local Similarity 58.8%; Pred. No. 5e-22;
Matches 50; Conservative 8; Mismatches 11; Indels 16; Gaps 1;
OY 1 TLVNGEFQVDCRSAGCGCOTSDNAAREALNASSKEEALQIIAAATPEKYLQFQHNLSNL 60
DB 102 TLWNGEFQVDCRSAGCGCQNDACAEALNASSKEEALQIIAAATPEKYLQFQHNLSNL 161
OY 61 DRI-----FCKTPE 69
DB 162 DRIAPLELVFVCPSSSSFDQVPE 186

```

```

RESULT 14
OQ20C0 ID OQ20C0 PRELIMINARY; PRT; 208 AA.
AC OQ20C0
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE C1 AND C4 GENES, CLONE AMG-1(B152),
DE PARTIAL AND COMPLETE CDS (FRAGMENT).
GN Tobacco leaf curl virus.
OS Viruses: ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=67762;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-AMG152
RA Ooi K., Ohshita S., Ishii I., Yahara T.;
RT "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL J. Plant Res. 110:247-257(1997).
DR EMBL: AB001256; BAA33995.1; -.
DR INTERPRO: IPRO01191; -.
DR PFAM: PF00799; Gemin1_AL1; 1.
DR PRINTS: PR00227; GEMCOATLVL.
DR NON_TER 208 208
FT NON_TER 208 208
SQ SEQUENCE 208 AA; 23486 MW; E301135F799C3DAD CRC64;

```

```

Query Match 69.5%; Score 251; DB 12; Length 208;
Best Local Similarity 57.6%; Pred. No. 8.7e-22;
Matches 49; Conservative 9; Mismatches 11; Indels 16; Gaps 1;
OY 1 TLVNGEFQVDCRSAGCGCOTSDNAAREALNASSKEEALQIIAAATPEKYLQFQHNLSNL 60
DB 102 TLWNGEFQVDCRSAGCGCQNDACAEALNASSKEEALQIIAAATPEKYLQFQHNLSNL 161
OY 61 DRI-----FCKTPE 69
DB 162 DRIAPLELVFVCPSSSSFDQVPE 186

```

```

RESULT 15
OQ20B8 ID OQ20B8 PRELIMINARY; PRT; 208 AA.
AC OQ20B8
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE C1 AND C4 GENES, CLONE AMG-2 (B154),
DE PARTIAL AND COMPLETE CDS (FRAGMENT).
GN Tobacco leaf curl virus.
OS Viruses: ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=67762;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-AMG154
RA Ooi K., Ohshita S., Ishii I., Yahara T.;
RT "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL J. Plant Res. 110:247-257(1997).
DR EMBL: AB001297; BAA33998.1; -.
DR INTERPRO: IPRO01191; -.
DR PFAM: PF00799; Gemin1_AL1; 1.
DR PRINTS: PR00227; GEMCOATLVL.
DR NON_TER 1 208
FT NON_TER 208 208
SQ SEQUENCE 208 AA; 23472 MW; 6290D00EF7C9956AA CRC64;

```

Query Match 59.5%; Score 251; DB 12; Length 208;  
Percent Similarity 57.6%; Strict Matches 11; Indels 16; Gaps 1;  
Matches 49; Conservative  
QY 1 TLVWGEFQVDCRSARGGCTSDNDAAEALNASSKKEALQITAAATPEKYLFOFHNLSNL 60  
DB 102 TLEWGTFTQIDGRSARGGCONANDAEALNASSKADALAITREKLPKDFIQYHLSNL 161  
QY 61 DRI-----FDKTPF 69  
DB 162 DRIFAPPLEVFCPPSSSFQVPE 186

Search completed: February 3, 2001, 02:22:51  
Job time: 1888 sec



```

AC P2758;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE ALI PROTEIN:
DE POTATO YELLOW MOSAIC VIRUS (Isolate Venezuela).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91311403; PubMed=1856690;
RA Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.:
RT "Nucleotide sequence of the infectious cloned DNA components of
RL J. Gen. Virol. 72:1515-1520(1991).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: D00940; BAA00782.1; -.
DR PIR: J00364; OOCVPT.
DR INTERPRO: IPRO011301; -.
DR PFAM: PF007799; Gemin_A1; 1.
DR PRINTS: PR00227; GEMCOATL1.
DR ATPTS: PR00228; GEMCOATCLVL1.
KW ATP-binding.
FT NP_BIND 222 229 ATP (POTENTIAL).
SQ SEQUENCE 361 AA; 40850 MW; 5627A33BF1264393 CRC64;

Query Match 69.0%; Score 249; DB 1; Length 361;
Best Local Similarity 68.1%; Pred. No. 1,1e-22;
Matches 47; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDSRGSGCOTSDNDAAEALNASSKEALQITAAAIPEKYLFOFHNSNL 60
DB 110 TVEWGQFQIDGSRGQGVNDAAENLNGSTGKAMKILKEAPKLFQFHNSNL 169
QY 61 DRIFDKPTPE 69
DB 170 DRIFDKAPE 178

RESULT 3
ALI_PROT 4
ID VALI CLVW STANDARD: PRT: 358 AA.
AC P14982;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DE ALI PROTEIN (40.4 KDA PROTEIN).
GN ACI.
OC Cassava latent virus (strain Nigerian).
CC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90174930; PubMed=2308831;
RA Morris B., Coates L., Lowe S., Richardson K., Eddy P.:
RT "Nucleotide sequence of the infectious cloned DNA components of
RT African cassava mosaic virus (Nigerian strain)".
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: X17095; CAA34953.1; -.
DR PIR: S07594; S07594.
DR INTERPRO: IPRO01191; -.
DR INTERPRO: IPRO011301; -.
DR PFAM: PF007799; Gemin_A1; 1.
DR PRINTS: PR00227; GEMCOATL1.
DR ATPTS: PR00228; GEMCOATCLVL1.
KW ATP-binding.
FT NP_BIND 220 227
SQ SEQUENCE 358 AA; 40435 MW; 1DB16B80CB2D5E2C CRC64;

Query Match 66.5%; Score 240; DB 1; Length 358;
Best Local Similarity 61.4%; Pred. No. 1,3e-21;
Matches 43; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDSRGSGCOTSDNDAAEALNASSKEALQITAAAIPEKYLFOFHNSNL 60
DB 109 TVEWGQFQIDGSRGQGVNDAAENLNGSTGKAMKILKEAPKLFQFHNSNL 168
QY 61 DRIFDKPTPE 70
DB 169 DRIFQEPAP 178

```

```

CC or send an email to license@isb-sib.ch).
CC EMBL: J02057; -. NOT_ANNOTATED_CDS.
DR INTERPRO: IPRO011301; -.
DR PFAM: PF007799; Gemin_A1; 1.
DR PRINTS: PR00227; GEMCOATL1.
DR ATPTS: PR00228; GEMCOATCLVL1.
KW ATP-binding.
FT NP_BIND 220 227 ATP (POTENTIAL).
SQ SEQUENCE 358 AA; 40346 MW; ED173E753E92D69 CRC64;

Query Match 66.5%; Score 240; DB 1; Length 358;
Best Local Similarity 61.4%; Pred. No. 1,3e-21;
Matches 43; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDSRGSGCOTSDNDAAEALNASSKEALQITAAAIPEKYLFOFHNSNL 60
DB 109 TVEWGQFQIDGSRGQGVNDAAENLNGSTGKAMKILKEAPKLFQFHNSNL 168
QY 61 DRIFDKPTPE 70
DB 169 DRIFQEPAP 178

RESULT 4
ALI_PROT 4
ID VALI CLVW STANDARD: PRT: 358 AA.
AC P14972;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DE ALI PROTEIN (40.4 KDA PROTEIN).
GN ACI.
OC Cassava latent virus (strain Nigerian).
CC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90174930; PubMed=2308831;
RA Morris B., Coates L., Lowe S., Richardson K., Eddy P.:
RT "Nucleotide sequence of the infectious cloned DNA components of
RT African cassava mosaic virus (Nigerian strain)".
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: X17095; CAA34953.1; -.
DR PIR: S07594; S07594.
DR INTERPRO: IPRO01191; -.
DR INTERPRO: IPRO011301; -.
DR PFAM: PF007799; Gemin_A1; 1.
DR PRINTS: PR00227; GEMCOATL1.
DR ATPTS: PR00228; GEMCOATCLVL1.
KW ATP-binding.
FT NP_BIND 220 227 ATP (POTENTIAL).
SQ SEQUENCE 358 AA; 40435 MW; 1DB16B80CB2D5E2C CRC64;

Query Match 66.5%; Score 240; DB 1; Length 358;
Best Local Similarity 61.4%; Pred. No. 1,3e-21;
Matches 43; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDSRGSGCOTSDNDAAEALNASSKEALQITAAAIPEKYLFOFHNSNL 60
DB 109 TVEWGQFQIDGSRGQGVNDAAENLNGSTGKAMKILKEAPKLFQFHNSNL 168

```

```

QY 61 DRIFDKTPEP 70
DB 169 DRIFQPPAP 178

RESULT 5
VALL TYLCU STANDARD; PRT; 362 AA.
AC P36279;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE ALL PROTEIN (C1 PROTEIN).
GN Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94256836; PubMed=8198442;
HA DRYL B. Hidalgo J., Accotto G., Moriones E.,
RT "Nucleotide sequence and genome organization of tomato leaf curl
geminivirus.";
J. Gen. Virol. 74:147-151(1993).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to licensed@isb-sib.ch).
DB EMBL: S53251; CAB30888.1;
DR PFAM: PF00799; Gemin1_AL1;
DR INTERPRO: IPR001301;
DR PRINTS: PR00227; GEMCONTALL1.
KW ATP-Binding. 221 228 ATP (POTENTIAL).
FT NP_BIND 221 228 ATP (POTENTIAL).
SQ SEQUENCE 362 AA; 41197 MW; 343E7E4547040598 CRC64;

Query Match 64.5%; Score 233; DB 1; Length 362;
Best Local Similarity 54.1%; Pred. No. 9.2e-21;
Matches 46; Conservative 14; Mismatches 14; Indels 16; Gaps 1;

QY 1 TLVWGFEQVGRSARGGCTSDNAAEALNASSKEALQITAAAPKYLQFPHNLNSLD 60
DB 110 TLEWGFQIDGRSAGGQSDANDAYAKAINTGSALNVLRELAPKDYVQLPHNLNSLD 169

QY 61 DRIFTPLVVPFSSSFORVPE 194
DB 170 DRIFTPLVVPFSSSFORVPE 194

RESULT 6
VALL TYLCU STANDARD; PRT; 359 AA.
AC P36609;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE ALL PROTEIN (C1 PROTEIN).
GN Tomato yellow leaf curl virus (strain Murcia) (TYLCV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94256836; PubMed=8198442;

Noris E., Hidalgo E., Accotto G., Moriones E.;
"High similarity among the tomato yellow leaf curl virus isolates
from the west Mediterranean basin: the nucleotide sequence of an
infectious clone from Spain.";
Arch. Virol. 135:165-170(1994).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to licensed@isb-sib.ch).
DB EMBL: Z25751; CAB1026.1;
DR PIR: S39211; S39211.
DR INTERPRO: IPR001191;
DR PFAM: PF00799; Gemin1_AL1;
DR PRINTS: PR00227; GEMCONTALL1.
KW ATP-Binding. 221 228 ATP (POTENTIAL).
FT NP_BIND 221 228 ATP (POTENTIAL).
SQ SEQUENCE 359 AA; 41065 MW; 2D170A51EF80A3EC CRC64;

Query Match 63.4%; Score 229; DB 1; Length 359;
Best Local Similarity 60.9%; Pred. No. 2.8e-20;
Matches 42; Conservative 8; Mismatches 19; Indels 0; Gaps 0;

QY 2 LVWGFEQVGRSARGGCTSDNAAEALNASSKEALQITAAAPKYLQFPHNLNSLD 61
DB 111 LEMWTFQIDGRSAGGQSDANDAYAKAINTGSKSEALDVIKELAPROYILPHNLNSLD 170

QY 62 RIFDKTPEP 70
DB 171 RIFQVPPAP 179

RESULT 7
VALL PHUV STANDARD; PRT; 349 AA.
AC Q06953;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE ALL PROTEIN.
GN All.
OS Pepper huasteco virus (PHV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94015007; PubMed=8409944;
HA Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,
RA Rivera-Bustamante R.F.;
RT "Complete nucleotide sequence of pepper huasteco virus: analysis and
comparison with bipartite geminiviruses.";
J. Gen. Virol. 74:2225-2231(1993).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to licensed@isb-sib.ch).
DB EMBL: X70418; CAB49556.1;
DR PIR: S31875; S31875.
DR PRINTS: JQ2300; JQ2300.
DR INTERPRO: IPR001191;

```

```

DR INTERPRO: IPR001301; ..
DR PFAM: PF00799; Gemini_ALI; 1.
DR PRINTS: PR00227; GEMCOATL1.
DR PROSITE: PS00228; GEMCOATLVL.
KW ATP-binding 221 228 ATP (BY SIMILARITY).
FT NP-BIND 221 228
SQ SEQUENCE 349 AA; 39722 MW; D5E4E76C5D6370E4 CRC64;

Query Match 61.5%; Score 222; DB 1; Length 349;
Best Local Similarity 60.0%; Pred. No. 1.9e-19;
Matches 42; Conservative 11; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLVGEQVDGSRAGCGQTNDRAAEALNASSKEALQIIAAAIPEKYLQFHNLSNL 60
   I: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 110 TVMEGEQVDGSRAGCGQTNDRAAEALNASSKEALQIIAAAIPEKYLQFHNLSNL 169
   I: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||

QY 61 DRIFQKTPPE 70
   I: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 170 NRIFQTPPE 179

RESULT 8
VAL1_TYLCM STANDARD; PRT; 359 AA.
AC P27260; 1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE ALL PROTEIN (40.2 KDA PROTEIN).
DE ALL PROTEIN (C1 PROTEIN).
GN CL.
OS Tomato yellow leaf curl virus (strain Marmande) (TYLCV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RS SEQUENCE FROM N.A.
RX MEDLINE-91107660; PubMed-1840676;
KW Kheyl-Hour A.; Bendabmane M.; Metzelt V.; Accotto G.P.; Crespi S.;
RA Gronenborn B.;
RT "Tomato yellow leaf curl virus from Sardinia is a
RT whitefly-transmitted monopartite geminivirus.";
RL Nucleic Acids Res. 19:6763-6769(1991).
CC -- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DB EMBL: X61153; CAA43466.1; ..
DR PIR: S22593; S22593.
DR INTERPRO: IPR001191; ..
DR INTERPRO: IPR001301; ..
DR PFAM: PF00799; Gemini_ALI; 1.
DR PRINTS: PR00227; GEMCOATL1.
DR PROSITE: PS00228; GEMCOATLVL.
KW ATP-binding 220 227 ATP (POTENTIAL).
FT NP-BIND 220 227
SQ SEQUENCE 359 AA; 40733 MW; 9717B4AC93EAF CRC64;

```

```

Query Match 61.5%; Score 222; DB 1; Length 359;
Best Local Similarity 58.0%; Pred. No. 1.9e-19;
Matches 40; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

QY 2 LVWGEQVDGSRAGCGQTNDRAAEALNASSKEALQIIAAAIPEKYLQFHNLSNL 61
   I: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 111 LVWGEQVDGSRAGCGQTNDRAAEALNASSKEALQIIAAAIPEKYLQFHNLSNL 170
   I: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||

QY 62 DRIFQKTPPE 70
   I: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||

```

```

Db 171 KVFQVPEAP 179

RESULT 9
VAL1_ABMV STANDARD; PRT; 353 AA.
ID VAL1_ABMV
AC P05175;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DE ALL PROTEIN (40.2 KDA PROTEIN).
DE ALL PROTEIN (40.2 KDA PROTEIN).
GN ACL. Golden mosaic virus
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RS SEQUENCE FROM N.A.
RA Howarth A.J.; Caton J.; Bossert M.; Goodman R.M.;
RT "Nucleotide sequence of bean golden mosaic virus and a model for gene
RT regulation in geminiviruses.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).
CC -- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DB EMBL: M10070; AAA46318.1; ..
DR INTERPRO: IPR001191; ..
DR INTERPRO: IPR001301; ..
DR PFAM: PF00799; Gemini_ALI; 1.
DR PRINTS: PR00227; GEMCOATL1.
DR PROSITE: PS00228; GEMCOATLVL.
KW ATP-binding 222 229 ATP (POTENTIAL).
FT NP-BIND 222 229
SQ SEQUENCE 353 AA; 40190 MW; 80FA79DF6029A34 CRC64;

```

```

Query Match 60.1%; Score 217; DB 1; Length 353;
Best Local Similarity 60.0%; Pred. No. 7.7e-18;
Matches 42; Conservative 11; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLVGEQVDGSRAGCGQTNDRAAEALNASSKEALQIIAAAIPEKYLQFHNLSNL 60
   I: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 110 TVMEGEQVDGSRAGCGQTNDRAAEALNASSKEALQIIAAAIPEKYLQFHNLSNL 169
   I: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||

QY 61 DRIFQKTPPE 70
   I: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 170 NRIFQVPEP 179

```

```

RESULT 10
VAL1_ABMV STANDARD; PRT; 355 AA.
ID VAL1_ABMV
AC P21947; 1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DE ALL PROTEIN.
DE ALL PROTEIN.
GN ACL.
OS Abutilon mosaic virus (isolate West India).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RS SEQUENCE FROM N.A.
RX Fritschmuth T.; Zimmatt G.; Jeske H.;
RT "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic
RT as well as eukaryotic features.";
RL Virology 178:461-468(1990).
CC -- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X15983; ; NOT_ANNOTATED_CDS.
DR PIR: A36214; OCVW1.
DR INTERPRO: IPR001191; .
DR PFAM: PF00759; Gemini_A1; 1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00228; GEMCOATCLVL1.
DR ATP-binding. 221 228 ATP (POTENTIAL).
FT NP-BIND 355 AA; 40257 MW; 16A2CANBA65251E95 CRC64;
SQ SEQUENCE 355 AA; 40257 MW; 16A2CANBA65251E95 CRC64;

Query Match 59.8%; Score 216; DR 1; Length 355;
Best Local Similarity 58.6%; Pred. No. 1e-18;
Matches 41; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

Qy 1 TLVWGFQVDRSARGCQTSNDAAAEALNASSKEEAQIIAAATPEKYLQFPHLNSNL 60
Db 110 TIENGFDQDRSANGQQTANSTAKALNMTSDQALQLEKDPKDFVLQHHNLLANA 169

Qy 61 DRIFDKTEP 70
Db 170 ERIFAKPEP 179

RESULT 11
VALI_TYLCV STANDARD; PRT; 358 AA.
AC P14951;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE AL1 PROTEIN (40.6 KDA PROTEIN).
OS Beet curly top virus (BCTV).
CC Virus; ssDNA viruses; Geminiviridae; Curtovirus.
RN [1]
RP SEQUENCE FROM N.A.
RA Stanley J., Markham P.G., Callis R.J., Plummer M.S.;
RT "The nucleotide sequence of an infectious clone of the geminivirus
RT beet curly top virus."
RL EMBL J. 5:1761-1767(1986).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X04144; ; NOT_ANNOTATED_CDS.
DR INTERPRO: IPR001191; .
DR INTERPRO: IPR001301; .
DR PFAM: PF00759; Gemini_A1; 1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00228; GEMCOATCLVL1.
DR ATP-binding. 222 228 ATP (POTENTIAL).
FT NP-BIND 358 AA; 40889 MW; 39AA5FEC30BC9C333 CRC64;
SQ SEQUENCE 358 AA; 40889 MW; 39AA5FEC30BC9C333 CRC64;

Query Match 59.0%; Score 213; DB 1; Length 358;
Best Local Similarity 55.7%; Pred. No. 2.4e-18;

```

```

Matches 39; Conservative 14; Mismatches 17; Indels 0; Gaps 0;

Qy 1 TLVWGFQVDRSARGCQTSNDAAAEALNASSKEEAQIIAAATPEKYLQFPHLNSNL 60
Db 110 TIENGFDQDRSANGQQTANSTAKALNMTSDQALQLEKDPKDFVLQHHNLLANA 169

Qy 61 DRIFDKTEP 70
Db 170 QKIFQRPDP 179

RESULT 12
VALI_TYLCV STANDARD; PRT; 361 AA.
AC O06657;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN.
RN [1]
RP SEQUENCE FROM N.A.
RA Abouzid A.M., Polston J.E.; Hiebert E.;
RT "The nucleotide sequence of tomato mottle virus, a new geminivirus
RT isolated from tomatoes in Florida."
RL EMBL J. 23:100-105(1994).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: L14460; AAC32414.1; .
DR PIR: JQ1870; JQ1870.
DR INTERPRO: IPR001191; .
DR INTERPRO: IPR001301; .
DR PFAM: PF00759; Gemini_A1; 1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00228; GEMCOATCLVL1.
DR NP-binding. 222 229 ATP (BY SIMILARITY).
FT NP-BIND 361 AA; 40516 MW; R138B65CEBAC6950 CRC64;
SQ SEQUENCE 361 AA; 40516 MW; R138B65CEBAC6950 CRC64;

Query Match 58.2%; Score 210; DB 1; Length 361;
Best Local Similarity 54.3%; Pred. No. 5.5e-18;
Matches 38; Conservative 15; Mismatches 17; Indels 0; Gaps 0;

Qy 1 TLVWGFQVDRSARGCQTSNDAAAEALNASSKEEAQIIAAATPEKYLQFPHLNSNL 60
Db 110 TIENGFDQDRSANGQQTANSTAKALNMTSDQALQLEKDPKDFVLQHHNLSNL 169

Qy 61 DRIFDKTEP 70
Db 170 ERIFAKPEP 179

RESULT 13
VALI_TYLCV STANDARD; PRT; 357 AA.
AC O06657;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN (C1 PROTEIN).
RN [1]
RP SEQUENCE FROM N.A.
RA Abouzid A.M., Polston J.E.; Hiebert E.;
RT "The nucleotide sequence of tomato mottle virus, a new geminivirus
RT isolated from tomatoes in Florida."
RL EMBL J. 23:100-105(1994).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X04144; ; NOT_ANNOTATED_CDS.
DR INTERPRO: IPR001191; .
DR INTERPRO: IPR001301; .
DR PFAM: PF00759; Gemini_A1; 1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00228; GEMCOATCLVL1.
DR ATP-binding. 222 228 ATP (POTENTIAL).
FT NP-BIND 357 AA; 40889 MW; 39AA5FEC30BC9C333 CRC64;
SQ SEQUENCE 357 AA; 40889 MW; 39AA5FEC30BC9C333 CRC64;

Query Match 59.0%; Score 213; DB 1; Length 357;
Best Local Similarity 55.7%; Pred. No. 2.4e-18;

```





Oy 22 NDAABALNNS-----SKEALQIIAAAIPEKYLQFHNLSNLD 62  
Db 93 NEAKAEKJNTPPEFKAMPEYKKQALVETPAKQAEVKKIQLPEREMODFY--NANKDQ 150  
Oy 63 JFDK 56  
Db 151 LEVK 154

Search completed: February 3, 2001, 02:24:26  
Job time: 638 sec

**THIS PAGE BLANK (USPTO)**

Genome version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:17:38 ; Search time 118.74 seconds  
(without alignments)  
40,029 Million cell updates/sec

Title: US-09-289-346a-4  
Perfect score: 361  
Sequence: 1 TLWGEFQVDSRGSGCOT.....FQPHINLSNLDIFDKTPEP 70

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues  
Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR.66:\*  
1: p1r1.\*  
2: p1r2.\*  
3: p1r3.\*  
4: p1r4.\*

Pred. No. is the number of results predicted by chance to have a score greater than the actual score being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	346	95.8	361	1 QOCVLI	Al1 protein - toma
2	249	69.0	361	1 QOCVPT	Al1 protein - toma
3	240	66.5	358	2 J07594	hypothetical prote
4	233	64.5	362	1 J01887	Al1 protein - toma
5	229	63.4	359	2 S92111	gene C1 protein -
6	222	61.5	349	2 J02300	replicase - pepper
7	222	61.5	349	2 S31875	Al1 protein - pepp
8	219	60.7	359	2 J01887	hypothetical prote
9	219	60.7	359	2 S32325	hypothetical prote
10	216	59.8	351	2 J02327	Al1 protein - Ind
11	216	59.8	355	1 QOCVW1	AVI protein - abut
12	213	59.0	385	2 S28360	Al1 protein - beet
13	210	58.2	358	1 J01870	Al1 protein - toma
14	200	55.4	357	1 QOCVC1	replication-associ
15	199	55.1	360	2 S59885	Al1 protein (clone
16	199	55.1	360	2 S59885	Al1 protein (clone
17	128	35.5	347	1 S4CVS1	FC gamma (19g) ac
18	72	19.9	587	2 JCL419	hypothetical prote
19	63.5	17.6	584	2 T19061	hypothetical prote
20	62.5	17.3	1006	2 S74952	probable peptidyl-
21	61.5	17.0	299	2 B71967	glycerolaldehyde-3-p
22	61.5	17.0	335	2 JSD164	probable phosphos
23	61.5	17.0	481	2 A70091	protein kinase sub
24	61.5	17.0	1293	2 T74915	protein kinase sub
25	61.5	17.0	1293	2 T74915	protein kinase sub
26	61.5	17.0	1293	2 T74915	protein kinase sub
27	60.5	16.8	159	2 C71838	NADH dehydrogenase
28	60.5	16.8	159	2 E64677	NADH dehydrogenase
29	60.5	16.8	334	1 DRS5GF	glycerolaldehyde-3-p

30 60 16.6 767 2 F71479 hypothetical prote  
31 60 16.6 1008 2 T41244 SEC14 protein homo  
32 59 16.3 160 2 G82060 hypothetical prote  
33 59 16.3 1502 1 K8YH1 C1C1/CFP3 transci  
34 58.5 16.2 357 2 T02246 hypothetical prote  
35 58.5 16.2 357 2 T02246 hypothetical prote  
36 58.5 16.2 769 2 F81742 conserved hypotet  
37 58.5 16.2 2135 2 T14602 variant-specific s  
38 58 16.1 555 2 C45868 glycerol-3-phospha  
39 57.5 15.9 858 2 C64577 alkali homolog - He  
40 57.5 15.9 1409 2 S74916 deoxyguanosine kin  
41 57 15.8 364 2 S75068 deoxyguanosine kin  
42 57 15.8 364 2 S75068 deoxyguanosine kin  
43 57 15.8 604 2 T19682 hypothetical prote  
44 57 15.8 771 2 B38252 granulocyte colony  
45 57 15.8 783 2 JH0329 granulocyte colony

## ALIGNMENTS

## RESULT 1

QOCVLI  
Al1 protein - tomato golden mosaic virus  
C:Species: tomato golden mosaic virus  
A:Note: host Nicotiana sp. (tobacco)  
C:Date: 28-Aug-1995 #sequence\_revision 28-Aug-1995 #text\_change 08-Apr-1994  
C:Accession: A04170 Steln. V.E.; Coutis, R.H.A.; Buck, K.W.  
C:Hamilton, 2197-2205, 1994  
A:Title: Complete nucleotide sequence of the infectious cloned DNA components of toma  
PMO, 2197-2205, 1994  
A:Reference number: A04163  
A:Accession: A04170  
A:Molecule type: DNA  
A:Residues: 1-352 <HAM>  
C:Comment: The genome consists of two circular, single-stranded DNA components, DNA A  
C:Genetics:  
A:Map position: segment A  
C:Superfamily: tomato golden mosaic virus Al1 protein

Query Match 95.8%; Score 346; DB 1; Length 352;  
Best local similarity 95.78; Pred. No. 2.6e-33;  
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
OY 1 TLWGEFQVDSRGSGCOTSDAAAEALNASSKEEALQIIIAAIPKYLTPQPHINLSNL 60  
DB 111 TLWGEFQVDSRGSGCOTSDAAAEALNASSKEEALQIIIAAIPKYLTPQPHINLSNL 170  
OY 61 DRIFDKTPEP 70  
DB 171 DRIFDKTPEP 180  
RESULT 2  
QOCVPT  
Al1 protein - potato yellow mosaic virus (isolate Venezuela)  
C:Species: potato yellow mosaic virus  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Jun-2000  
C:Accession: J00364  
C:Hamilton, 2197-2205, 1994  
A:Title: The nucleotide sequence of the infectious cloned DNA components of potato ye  
A:Reference number: J00362; MUID:91131403  
A:Accession: J00364  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-352 <HAM>  
C:Comment: The genome consists of two circular, single-stranded DNA components, DNA A  
C:Genetics:  
A:Map position: segment A  
C:Superfamily: tomato golden mosaic virus Al1 protein

```

Query Match          69.0%; Score 249; DB 1; Length 361;
Best Local Similarity 61.1%; Pred. No. 8.3e-23;
Matches 47; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSGAGCGCOTSDNDAARALNASSKEALQITAAAIPEKYLQFHNLSNL 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 110 TLVWGEFQVDRSGAGCGCOTVNDAAARALNASSKEALQITAAAIPEKYLQFHNLSNL 169
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 DRIFDKTPEP 69
Db 170 DRIFDKTPEP 178

RESULT 3
S07594
hypothetical protein, 40.4k - cassava latent virus (Nigerian isolate)
C:Species: cassava latent virus
C:Dates: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C:Accession: S07594
R:Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.
Nucleic Acids Res. 18, 197-198, 1990
A:Title: Nucleotide sequence of the infectious cloned DNA components of African cassava
A:Reference number: S07590; MUID:90174930
A:Accession: S07594
A:Status: translation not shown
A:Map position: segment DNA
A:Residues: 1-358 <ORF>
A:Cross-references: EMBL:X17095; NID:g59371; PIDN:CA34953.1; PID:g59376
C:Genetics:
C:Map position: segment DNA
C:Superfamily: tomato golden mosaic virus All protein

Query Match          66.5%; Score 240; DB 2; Length 358;
Best Local Similarity 61.4%; Pred. No. 9.5e-21;
Matches 43; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSGAGCGCOTSDNDAARALNASSKEALQITAAAIPEKYLQFHNLSNL 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 109 TLVWGEFQVDRSGAGCGCOTSDNDAARALNASSKEALQITAAAIPEKYLQFHNLSNL 168
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 DRIFDKTPEP 70
Db 169 DRIFDKTPEP 178

RESULT 4
J01887
Nucleic acid - tomato yellow leaf curl virus (strain Australia)
N:Alternate names: CL protein
C:Species: tomato yellow leaf curl virus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C:Accession: J01887
R:Dry, I.B.; Rigden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.
J. Gen. Virol. 74, 147-151, 1993
A:Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.
A:Reference number: J01887; MUID:9139778
A:Accession: J01887
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-362 <DR>
A:Cross-references: GB:S53251
C:Superfamily: tomato golden mosaic virus All protein

Query Match          64.5%; Score 233; DB 1; Length 362;
Best Local Similarity 54.1%; Pred. No. 6.5e-20;
Matches 46; Conservative 9; Mismatches 14; Indels 16; Gaps 1;

QY 1 TLVWGEFQVDRSGAGCGCOTSDNDAARALNASSKEALQITAAAIPEKYLQFHNLSNL 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

---

```

Db 110 TLVWGEFQVDRSGAGCGCOTSDNDAARALNASSKEALQITAAAIPEKYLQFHNLSNL 169
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 DRIFDKTPEP 69
Db 170 DRIFDKTPEP 178

RESULT 5
S39211
gene C1 protein - tomato yellow leaf curl virus
C:Species: tomato yellow leaf curl virus
C:Dates: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
C:Accession: S39211
R:Norris, E.; Hidalgo, E.; Accotto, G.; Moriones, E.
Submitted to the EMBL Data Library, August 1993
A:Description: High similarity among the tomato yellow leaf curl virus isolates from
A:Reference number: S39209
A:Accession: S39211
A:Status: preliminary
A:Molecule type: <ORF>
A:Residues: 1-359 <ORF>
A:Cross-references: EMBL:Z25751; NID:g433655; PIDN:CAAB1026.1; PID:g433658
C:Superfamily: tomato golden mosaic virus All protein

Query Match          63.4%; Score 229; DB 2; Length 359;
Best Local Similarity 60.9%; Pred. No. 1.9e-19;
Matches 42; Conservative 6; Mismatches 19; Indels 0; Gaps 0;

QY 2 LVWGEFQVDRSGAGCGCOTSDNDAARALNASSKEALQITAAAIPEKYLQFHNLSNL 61
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 111 LVWGEFQVDRSGAGCGCOTANDAYAKAINAGSKSEALDVKEIAPRDYILHFHNLSNL 170
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 62 DRIFDKTPEP 70
Db 171 DRIFDKTPEP 179

RESULT 6
J02300
replicase - pepper huasteco virus (component A)
N:Alternate names: ORF All protein
C:Species: pepper huasteco virus
C:Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Sep-1999
C:Accession: J02300
R:Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante,
J. Gen. Virol. 74, 2225-2231, 1993
A:Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b
A:Reference number: J02299; MUID:94015007
A:Accession: J02300
A:Molecule type: RNA
A:Residues: 1-919 <ORF>
A:Cross-references: GB:X70418; NID:g61023; PIDN:CAAA9856.1; PID:g61025
C:Superfamily: tomato golden mosaic virus All protein

Query Match          61.5%; Score 222; DB 2; Length 349;
Best Local Similarity 60.0%; Pred. No. 1.3e-18;
Matches 42; Conservative 11; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSGAGCGCOTSDNDAARALNASSKEALQITAAAIPEKYLQFHNLSNL 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 110 TLVWGEFQVDRSGAGCGCOTANDYAKAINAGSKSEALDVKEIAPRDYILHFHNLSNL 169
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 DRIFDKTPEP 70
Db 170 DRIFDKTPEP 179

RESULT 7
S31875
All protein - pepper rizado amarillo virus
C:Species: pepper rizado amarillo virus

```

C>Date: 22-Nov-1993 #sequence\_revision 26-May-1995 #text\_change 20-Sep-1999

A:Accession: S31875

R:Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Jimenez, B.; Herrera-Estrella, L.; Rivera-Bu

Submitted to the EMBL Data Library, February 1993

A:Description: Complete nucleotide sequence of Pepper huasteco virus: analysis and compa

A:Reference: S31875

A:Accession: S31875

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-349 <TOR>

A:Cross-references: EMBL:X70418; NID:g61023; PIDN:CAA9856.1; PID:g61025

A:Note: The source is designated as pepper huasteco virus

C:Superfamily: tomato golden mosaic virus ALL protein

Query Match 61.5%; Score 222; DB 2; Length 349;

Best Local Similarity 60.0%; Pred. No. 1.3e-18;

Matches 42; Conservative 11; Mismatches 17; Indels 0; Gaps 0;

Oy 1 TLVGEFQVDRSGRGCGTSDNDAAEALNASSKEEALQITAAAIPEKYLQFPHNLSNL 60

Db 111 TLVGEFQVDRSGRGCGTSDNDAAEALNASSKEEALQITAAAIPEKYLQFPHNLSNL 60

Oy 61 DRIFDKTPPEP 70

Db 170 NRIFQTPPEP 179

RESULT 8

S22593

hypothetical protein C4 - tomato yellow leaf curl virus

C:Species: Tomato yellow leaf curl virus

C>Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 20-Sep-1999

A:Accession: S22593

R:Kheyr-Pour, A.; Bendahmane, M.; Matzelt, V.; Accotto, G.P.; Crespi, S.; Gronenborn, B.

A:Title: Tomato yellow leaf curl virus (TYLCV) from Sardinia is a whitefly-transmitted monoparti

A:Reference number: S22588; MUID:92107660

A:Cross-references: EMBL:X1153; NID:g62211; PIDN:CAA43466.1; PID:g62217

A:Note: The source is designated as TYLCV isolate from Sardinia, August 1991

C:Superfamily: tomato golden mosaic virus ALL protein

Query Match 61.5%; Score 222; DB 2; Length 359;

Best Local Similarity 58.0%; Pred. No. 1.3e-18;

Matches 40; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

Oy 2 LVWGEFQVDRSGRGCGTSDNDAAEALNASSKEEALQITAAAIPEKYLQFPHNLSNL 61

Db 111 LVWGEFQVDRSGRGCGTSDNDAAEALNASSKEEALQITAAAIPEKYLQFPHNLSNL 61

Oy 62 RIFDKTPPEP 70

Db 171 KVFOVPPAP 179

RESULT 9

S39235

gene C1 protein - tomato yellow leaf curl virus

C:Species: Tomato yellow leaf curl virus

C>Date: 05-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 08-Sep-1997

A:Accession: S39235

R:Crute, S.; Norris, E.; Valera, A.; Bosco, D.; Accotto, G.

A:Title: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity.

A:Reference number: S39233

A:Cross-references: S39235

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-359 <CRE>

A:Cross-references: EMBL:Z28390; NID:g1041671; PID:g1334964

C:Superfamily: tomato golden mosaic virus ALL protein

Query Match 60.7%; Score 219; DB 2; Length 359;

Best Local Similarity 58.0%; Pred. No. 2.9e-18;

Matches 40; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

Oy 2 LVWGEFQVDRSGRGCGTSDNDAAEALNASSKEEALQITAAAIPEKYLQFPHNLSNL 61

Db 111 LVWGEFQVDRSGRGCGTSDNDAAEALNASSKEEALQITAAAIPEKYLQFPHNLSNL 61

Oy 62 RIFDKTPPEP 70

Db 171 KVFOVPPAP 179

RESULT 10

JQ2327

protein - Indian casaca mosaic virus

C:Species: Indian casaca mosaic virus

C>Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 20-Sep-1999

A:Accession: JQ2327; S35883

R:Hong, Y.G.; Robinson, D.J.; Harrison, B.D.

A:Title: Nucleotide sequence evidence for the occurrence of three distinct whitefly-t

A:Reference number: JQ2326; MUID:94065970

A:Cross-references: EMBL:Z24758; NID:g395351; PIDN:CAA80891.1; PID:g584046

A:Note: The source is designated as Indian casaca mosaic virus ALL protein

C:Superfamily: tomato golden mosaic virus ALL protein

Query Match 59.8%; Score 216; DB 2; Length 351;

Best Local Similarity 61.2%; Pred. No. 6.5e-18;

Matches 41; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

Oy 4 WGEFQVDRSGRGCGTSDNDAAEALNASSKEEALQITAAAIPEKYLQFPHNLSNLDRT 63

Db 113 WGTFTQIDRSRGCGTSDNDAAEALNASSKEEALQITAAAIPEKYLQFPHNLSNLDRT 172

Oy 64 EDKTPPEP 70

Db 173 FTSPEPP 179

RESULT 11

QCCWI

protein - abutilon mosaic virus (isolate West India)

C:Species: Abutilon mosaic virus

C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 08-Apr-1994

A:Accession: A36214

R:Prischmuth, T.; Zimmat, G.; Jeske, H.

A:Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as

A:Reference number: A36214; MUID:91020984

A:Cross-references: A36214

A:Note: The source is designated as abutilon mosaic virus

C:Superfamily: tomato golden mosaic virus ALL protein

Query Match 59.8%; Score 216; DB 1; Length 355;

Best Local Similarity 58.6%; Pred. No. 6.6e-18;

Matches 41; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

Oy 1 TLVGEFQVDRSGRGCGTSDNDAAEALNASSKEEALQITAAAIPEKYLQFPHNLSNL 60

Db 110 TAAKGCFQDGRSARGGQTSNDAAALNASSKEALQITAAAIPEKYLQFHNLNSNL 169  
 QY 61 DRIFDKTPEP 70  
 C:Date 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 20-Sep-1999  
 C:Accession: J01870  
 Db 170 BRIFAKAPEP 179

## RESULT 12

S28360  
 A:Title: protein - beet curly top virus  
 A:Accession: S28360  
 C:Date 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 20-Sep-1999  
 C:Accession: J01870  
 R:Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.  
 EMBL J. 5, 1761-1767, 1986  
 A:Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly top  
 A:Reference number: S28360  
 A:Accession: S28360  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-385 <STA>  
 A:Cross-references: GR:M24597; EMBL:X04144; NID:g210678; PIDN:AAA42751.1; PID:g210679  
 C:Superfamily: tomato golden mosaic virus All protein

Query Match 50.0%; Score 213; DB 2; Length 385;  
 Best Local Similarity 55.4%; Pred. No. 5, 2e-16;  
 Matches 39; Conservative 14; Mismatches 17; Indels 0; Gaps 0;  
 QY 1 TLVNGCFQDGRSARGGQTSNDAAALNASSKEALQITAAAIPEKYLQFHNLNSNL 60  
 Db 137 TIVNGCFQDGRSARGGQTSNDAAALNASSKEALQITAAAIPEKYLQFHNLNSNL 196  
 QY 61 DRIFDKTPEP 70  
 Db 197 OKIFQRPDP 206

## RESULT 13

QJ01870  
 A:Title: protein - tomato mottle virus (isolate Florida)  
 A:Accession: QJ01870  
 C:Date 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 07-May-1999  
 C:Accession: J01870  
 R:Abouid, A.M.; Polston, J.E.; Hiebert, E.  
 J. Gen. Virol. 73, 3225-3229, 1992  
 A:Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated from  
 A:Reference number: JQ1869; MUID:93107858  
 A:Accession: JQ1870  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-358 <ABO>  
 A:Cross-references: GB:I14460  
 C:Genetics:  
 A:Map position: segment A  
 C:Superfamily: tomato golden mosaic virus All protein

Query Match 58.2%; Score 210; DB 1; Length 358;  
 Best Local Similarity 54.3%; Pred. No. 3, 4e-17;  
 Matches 38; Conservative 15; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLVNGCFQDGRSARGGQTSNDAAALNASSKEALQITAAAIPEKYLQFHNLNSNL 60  
 Db 107 TIVNGCFQDGRSARGGQTSNDAAALNASSKEALQITAAAIPEKYLQFHNLNSNL 166  
 QY 61 DRIFDKTPEP 70  
 Db 167 BRIFAKAPEP 176

## RESULT 14

QJ01870  
 A:Title: protein - tomato mottle virus (isolate Florida)  
 A:Accession: QJ01870  
 C:Date 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 07-May-1999  
 C:Accession: J01870  
 R:Abouid, A.M.; Polston, J.E.; Hiebert, E.  
 J. Gen. Virol. 73, 3225-3229, 1992  
 A:Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated from  
 A:Reference number: JQ1869; MUID:93107858  
 A:Accession: JQ1870  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-358 <ABO>  
 A:Cross-references: GB:I14460  
 C:Genetics:  
 A:Map position: segment A  
 C:Superfamily: tomato golden mosaic virus All protein

QJ01870  
 A:Title: protein - tomato yellow leaf curl virus  
 A:Accession: QJ01870  
 C:Date 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 23-Jul-1999  
 C:Accession: J01870  
 R:Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.  
 Virol. 185, 151-161, 1991  
 A:Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin  
 A:Reference number: A40779; MUID:92024070  
 A:Accession: A40779  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-357 <NAV>  
 A:Cross-references: GR:X15656; NID:g62204; PIDN:CAA33688.1; PID:g62207  
 C:Superfamily: tomato golden mosaic virus All protein

Query Match 55.4%; Score 200; DB 1; Length 357;  
 Best Local Similarity 63.9%; Pred. No. 5, 2e-16;  
 Matches 39; Conservative 9; Mismatches 13; Indels 0; Gaps 0;  
 QY 4 NGCFQDGRSARGGQTSNDAAALNASSKEALQITAAAIPEKYLQFHNLNSLDRI 63  
 Db 111 FGVSIDGRSARGGQTSNDAAALNASSKEALQITAAAIPEKYLQFHNLNSLDRI 170  
 QY 64 F 64  
 Db 171 P 171

## RESULT 15

S59885  
 A:Title: replication-associated protein C1 - tomato yellow leaf curl virus  
 A:Accession: S59885  
 C:Date 01-Mar-1996 #sequence\_revision 01-Mar-1996 #text\_change 20-Sep-1999  
 C:Accession: S59885  
 R:Hong, Y.; Harrison, B.D.  
 submitted to the EMBL Data Library, February 1995  
 A:Description: Nucleotide sequences from tomato leaf curl viruses from different coun  
 d geminiviruses.  
 A:Reference number: S58346  
 A:Accession: S59885  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-360 <HON>  
 A:Cross-references: EMBL:248182; NID:g944838; PIDN:CAA88229.1; PID:g974211  
 C:Superfamily: tomato golden mosaic virus All protein

Query Match 55.1%; Score 199; DB 2; Length 360;  
 Best Local Similarity 57.6%; Pred. No. 6, 1e-16;  
 Matches 38; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

QY 4 NGCFQDGRSARGGQTSNDAAALNASSKEALQITAAAIPEKYLQFHNLNSLDRI 63  
 Db 113 FGVSIDGRSARGGQTSNDAAALNASSKEALQITAAAIPEKYLQFHNLNSLDRI 172  
 QY 64 FDXTP 69  
 Db 173 FTFSAE 178

Search completed: February 3, 2001, 02:17:38  
 Job time: 2210 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:19:42 : Search time 109.09 Seconds  
(without alignments)  
11,523 Million cell updates/sec

Title: US-09-289-346a-4  
Pairwise score: 361  
Sequence: 1 TLWKEFOVDSRSGSQQT.....QFHLNLSNLDHFIDTEP 70

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/1/aa/5A.COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/aa/5B.COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/aa/6.COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/aa/PTCUS.COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result	Score	Query Match	Length	ID	Description
1	222	61.5	359	3	US-08-809-103B-2
2	222	61.5	359	3	US-08-809-103B-4
3	222	61.5	359	3	US-08-809-103B-6
4	222	61.5	359	3	US-08-809-103B-8
5	60.5	16.8	334	5	5290690-11
6	58.5	15.8	335	1	US-08-809-562-1
7	58.5	15.8	335	1	US-08-809-562-2
8	57	15.8	771	1	US-07-923-976-6
9	57	15.8	783	5	5422248-2
10	57	15.8	836	1	US-07-923-976-4
11	57	15.8	863	1	US-07-923-976-8
12	55	15.2	783	5	5231168-2
13	54.5	15.1	1713	3	US-08-600-982-24
14	54.5	15.1	1713	3	US-08-600-982-24
15	54.5	15.0	1271	1	US-08-276-919-10
16	54	15.0	271	1	US-08-776-088-13
17	54	15.0	271	4	PTC-US95-09145A-13
18	54	15.0	274	1	US-08-776-088-15
19	54	15.0	274	4	PTC-US95-09145A-15
20	54	15.0	325	1	US-08-276-919-13
21	54	15.0	325	1	US-08-276-919-13
22	54	15.0	325	1	US-08-776-088-18
23	54	15.0	325	1	US-08-776-088-18
24	54	15.0	325	4	PTC-US95-09145A-4
25	54	15.0	325	4	PTC-US95-09145A-4
26	54	15.0	354	1	US-08-276-919-18
27	54	15.0	354	1	US-08-276-919-2
28	54	15.0	354	1	US-08-776-088-6

Query Match 61.5% Score 222; DB 3; Length 359;  
Best Local Similarity 58.0% Pred. No. 2.6e-22;  
Matches 40; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

29 54 15.0 354 3 US-09-325-320-2  
30 54 15.0 354 4 PCT-US95-09145A-2  
31 54 15.0 354 4 PCT-US95-09145A-6  
32 53 14.7 297 3 US-09-027-013-3  
33 53 14.7 844 3 US-09-029-267-20  
34 53 14.7 844 3 US-09-029-267-20  
35 53 14.7 852 1 US-08-130-802A-59  
36 53 14.7 852 1 US-08-130-802A-67  
37 52.5 14.5 3898 2 US-08-876-991-2  
38 52.5 14.5 3898 2 US-09-059-853-2  
39 52 14.4 207 2 US-08-177-109A-60  
40 52 14.4 207 2 US-08-687-706-50  
41 52 14.4 482 1 US-07-792-885A-11  
42 52 14.4 482 1 US-07-792-885A-11  
43 52 14.4 482 1 US-08-142-439A-7  
44 52 14.4 482 2 US-08-869-477-7  
45 51.5 14.3 870 2 US-08-266-311-1

## ALIGNMENTS

RESULT 1  
US-08-809-103B-2  
: Sequence 2, Application US/08809103B  
: Patent No. 6133505  
: GENERAL INFORMATION:  
: ATTORNEY: GRONENHOFF, BEUNG  
: TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME  
: NUMBER OF SEQUENCES: 17  
: CORRESPONDENCE ADDRESS:  
: ADDRESSER: YOUNG & THOMPSON  
: STREET: 745 South 23rd Street  
: CITY: Arlington  
: STATE: Virginia  
: COUNTRY: U.S.A.  
: ZIP: 22202  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.30  
: CURRENT OFFICE ACTION NUMBER: US/08/809,103B  
: APPLICATION NUMBER: US/08/809,103B  
: FILING DATE: 17-MAR-1997  
: CLASSIFICATION: 800  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: 94.11040  
: FILING DATE: 15-SEP-1994  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: NO PCT/F95/01192  
: FILING DATE: 15-SEP-1995  
: ATTORNEY/AGENT INFORMATION:  
: NAME: PATCH, Andrew J.  
: REGISTRATION NUMBER: 32,925  
: REFERENCE/DOCKET NUMBER: USB94AL CNR TOM  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (703) 685-2297  
: TELEFAX: (703) 685-2273  
: TELEX: 248425 EMBON  
: INFORMATION FOR SEQ ID NO: 2:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 359 amino acids  
: TYPE: amino acid  
: TOPOLOGY: Linear  
: MOLECULE TYPE: protein  
US-08-809-103B-2

QY 2 LVWGFVQDGRSARGCQTSNDAAALMASKEEALQIIAAIPKYLQFPHNLNSLD 61  
Db 111 LEWTFQIDGRSARGCQQTANDAYAKAINAGSKQALDIKELAPRDYVLHFNHNSLD 170  
QY 62 RIFDKTPEP 70  
Db 171 KVFQVPPAP 179  
RESULT 2  
US-08-809-103b-4  
: Sequence 4, Application US/06809103B  
: Patent No. 6133505  
: GENERAL INFORMATION:  
: APPLICANT: GRONENBORN, Bruno  
: TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT  
: NUMBER OF SEQUENCES: 17  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: YOUNG & THOMPSON  
: STREET: 745 South 23rd Street  
: CITY: Arlington  
: STATE: Virginia  
: COUNTRY: U.S.A.  
: ZIP: 22202  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: FILING DATE: 17-MAR-1997  
: PRIORITY APPLICATION NUMBER: FR 94.11040  
: PRIOR APPLICATION DATA:  
: FILING DATE: 15-SEP-1994  
: APPLICATION NUMBER: WO PCT/FR95/01192  
: ATTORNEY/AGENT INFORMATION:  
: NAME: PATCH, Andrew J.  
: REGISTRATION NUMBER: 32,925  
: REFERENCE/DOCKET NUMBER: USB94AL CNR TOM  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (703) 521-2297  
: TELEFAX: 248425 EMBON  
: INFORMATION FOR SEQ ID NO: 4:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 359 amino acids  
: TYPE: amino acid  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
US-08-809-103b-6

Query Match 61.5%; Score 222; DB 3; Length 359;  
Best Local Similarity 58.0%; Pred. No. 2,6e-22;  
Matches 40; Conservative 10; Mismatches 19; Indels 0; Gaps 0;  
QY 2 LVWGFVQDGRSARGCQTSNDAAALMASKEEALQIIAAIPKYLQFPHNLNSLD 61  
Db 111 LEWTFQIDGRSARGCQQTANDAYAKAINAGSKQALDIKELAPRDYVLHFNHNSLD 170  
QY 62 RIFDKTPEP 70  
Db 171 KVFQVPPAP 179

RESULT 3  
US-08-809-103b-6

: Sequence 6, Application US/06809103B  
: Patent No. 6133505  
: GENERAL INFORMATION:  
: APPLICANT: GRONENBORN, Bruno  
: TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT  
: NUMBER OF SEQUENCES: 17  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: YOUNG & THOMPSON  
: STREET: 745 South 23rd Street  
: CITY: Arlington  
: STATE: Virginia  
: COUNTRY: U.S.A.  
: ZIP: 22202  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: FILING DATE: 17-MAR-1997  
: PRIORITY APPLICATION NUMBER: FR 94.11040  
: PRIOR APPLICATION DATA:  
: FILING DATE: 15-SEP-1994  
: APPLICATION NUMBER: WO PCT/FR95/01192  
: ATTORNEY/AGENT INFORMATION:  
: NAME: PATCH, Andrew J.  
: REGISTRATION NUMBER: 32,925  
: REFERENCE/DOCKET NUMBER: USB94AL CNR TOM  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (703) 521-2297  
: TELEFAX: 248425 EMBON  
: INFORMATION FOR SEQ ID NO: 6:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 359 amino acids  
: TYPE: amino acid  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
US-08-809-103b-6

Query Match 61.5%; Score 222; DB 3; Length 359;  
Best Local Similarity 58.0%; Pred. No. 2,6e-22;  
Matches 40; Conservative 10; Mismatches 19; Indels 0; Gaps 0;  
QY 2 LVWGFVQDGRSARGCQTSNDAAALMASKEEALQIIAAIPKYLQFPHNLNSLD 61  
Db 111 LEWTFQIDGRSARGCQQTANDAYAKAINAGSKQALDIKELAPRDYVLHFNHNSLD 170  
QY 62 RIFDKTPEP 70  
Db 171 KVFQVPPAP 179

RESULT 4  
US-08-809-103b-8  
: Sequence 8, Application US/06809103B  
: Patent No. 6133505  
: GENERAL INFORMATION:  
: APPLICANT: GRONENBORN, Bruno  
: TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT  
: NUMBER OF SEQUENCES: 17  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: YOUNG & THOMPSON  
: STREET: 745 South 23rd Street  
: CITY: Arlington  
: STATE: Virginia



```

; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; 615PIPTF100PPY disk
; COMPUTER: IBM PC
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/809,103B
; FILING DATE: 17-MAR-1997
; CLASSIFICATION: 800
; PRIORITY DATA:
; APPLICATION NUMBER: FR 94.11040
; FILING DATE: 15-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR95/01192
; ATTORNEY/AGENT INFORMATION:
; NAME: MATCH, Andrew J.
; REGISTRATION NUMBER: 2,925
; REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 521-2297
; TELEFAX: (703) 685-0573
; TELEX: 248425 PMRON
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; TYPE: amino acid
; TOPOLOGY: Linear
; MOLECULE TYPE: protein
; US-08-809-103B-8

Query Match 51.5%; Score 222; DB 3; Length 359;
Best Local Similarity 51.0%; Pred No 2.6;
Matches 40; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

Oy 2 LWKGEFVDVHSARGGQTSNDAAAEALNASSKKEALQITAAIPEKYLFOFHNLSNLD 61
DB 111 LEMQTFQIDHSARGGQGTANDAYAKAINAGSKSOALQVKEIAPROYVYLHFNLSNLD 170

Oy 62 RIFQKTPPEP 70
DB 171 KVFQEPAP 179

RESULT 5
5290690-11
; Patent No. 5290690
; INVENTOR: MONTI, LASTERS, IGWAGE, STANGSENS, PATRICK
; MATTHEWS; GASTON, WODAK, SHOSHAK, OUDY
; TITLE OF INVENTION: METHODS AND MEANS FOR CONTROLLING THE
; STABILITY OF PROTEINS
; NUMBER OF SEQUENCES: 22
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/398,706
; FILING DATE: 25-AUG-1989
; SEQ ID NO: 11
; LENGTH: 334
5290690-11

Query Match 16.8%; Score 60.5; DB 5; Length 334;
Best Local Similarity 28.4%; Pred. No. 2.6;
Matches 19; Conservative 11; Mismatches 28; Indels 9; Gaps 1;

Oy 2 LWKGEFVDVHSARGGQTSNDAAAEALNASSKKE-----EALQITAAIPEKYLFO 52
DB 82 LAWGEIGVDIVESTGRFTKDEAKHLKAGAKKVIISAPAKVNTVVMGVNQDKYDPK 141

Oy 53 FINLNS 59
; 1::11

```

```

DB 142 AHHVSN 148

RESULT 6
US-08-781-562-1
; Sequence 1: Application US/08781562
; Patent No. 5763589
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN MEMBRANE PROTEIN
; INVENTOR: HILLMAN, JENNIFER L.; GOLI, SURYA K.
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Inocyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq For Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,562
; FILING DATE: Herewith
; CLASSIFICATION: 500
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0181 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0655
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 356 amino acids
; TYPE: amino acid
; TOPOLOGY: single
; STABILITY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Consensus
; CLONE: Consensus
; US-08-781-562-1

Query Match 16.2%; Score 58.5; DB 1; Length 356;
Best Local Similarity 27.6%; Pred No 5.3;
Matches 16; Conservative 12; Mismatches 17; Indels 13; Gaps 1;

Oy 19 QTSNDAAAEALNASSKKEALQITAAI-----IPEKYLFOFHNLSNLDRI 63
DB 239 QAGEASAVLNKAKAKAEIRILAAUTQHNQNGDAASLTVAEQYVSFAFKLSKDSMTI 296

RESULT 7
US-08-419-652-6
; Sequence 6, Application US/08419652
; Patent No. 5831007
; GENERAL INFORMATION:
; APPLICANT: Chua, Anne O
; APPLICANT: Glick, Ulrich A
; TITLE OF INVENTION: THERAPEUTICALLY ACTIVE TETRALEUKIN-12 RECEPTOR
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley

```

```

Query Match      15.8% ; DB 2; Length 602;
Best Local Similarity   24.0%; Pred. No. 17;
Matches 12; Conservative    9; Mismatches 23; Indels    6; Caps    1;

QY       7 FVDGSRGSGCTDAAAEALNASSKEAALQIIIAATPKYLQFHNL 56
        1 : ||||| : : : ||| : | :
DB       37 FTLKFASFGNCOTGDGSILDCVPGKGSH-----CCIPKHLLYKNM 80

RESULT      8
US-07-923-976-6
: Sequence 6, Application US/07923976
: Patent No. 5574136
: GENERAL INFORMATION:
: APPLICANT: Nagata, Shigekazu
: APPLICANT: Fukunaga, Rikio
: TITLE OF INVENTION: DNA Encoding Granulocyte
: TITLE OF INVENTION: Colony-Stimulating Factor Receptor
: NUMBER OF SEQUENCES: 8
: CROSS-REFERENCE(S)
: ADDRESSES:
: STREET: P. O. Box 2266 Bads Station
: CITY: Arlington
: STATE: Virginia
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: SOFTWARE SYSTEM: PC-DOS/MS-DOS
: SOFTWARE RELEASE W1.0, Version #1.25
: CURRENT APPLICATION DATA
: APPLICATION NUMBER: US/07923-976
: FILING DATE: 19920922

```

```

: CLASSIFICATION: 435
: PRIOR APPLICATION NUMBER: JP 74539/1990
: FILING DATE: 23-MAR-1990
: PRIOR APPLICATION NUMBER: JP 176629/1990
: FILING DATE: 03-JUL-1990
: PCT/JP91/00375
: APPLICATION NUMBER: PCT/JP91/00375
: FILING DATE: 22-MAR-1991
: NAME: Hellwege, James W.
: REGISTRATION NUMBER: 28,808
: REFERENCE/DOCKET NUMBER: 514853
: TELEPHONE: 703-415-1500
: TELEFAX: 703-415-1508
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 771 amino acids
: TYPE: AMINO ACID
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-07-923-976-6

Query Match 15.8%; Score 57; DB 1; Length 771;
Best Local Similarity 24.0%; Pred. No. 24;
Matches 12; Conservative 9; Mismatches 23; Indels 6

QY 7 FQVDSRGAGCGQTSNDAAEAALNASSKEEAQITAAATPKYILFQFNL 56
DB 156 PTLKSFSGNCGTQGSILDCVPRQGSH-----CCIPRKHLLYQNM 199

RESULT 9
5422248-2
: Patent No. 5422248
: INVENTOR: CURTIS SMITH, CRAIG A.; LARSEN, ALF D.; SIMS, JOHN E.;
: BENSON, CURTIS
: TITLE OF INVENTION: DNA SEQUENCES ENCODING GRANULOCYTE-COLONY
: STIMULATING FACTOR RECEPTORS
: NUMBER OF SEQUENCES: 6
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/6,183
: FILING DATE: 15-JAN-1993
: PRIORITY CLAIM: YES
: APPLICATION NUMBER: 587,329
: FILING DATE: 24-SEP-1990
: APPLICATION NUMBER: 522,952
: FILING DATE: 03-APR-1990
: APPLICATION NUMBER: 416,306
: FILING DATE: 03-OCT-1989
: APPLICATION NUMBER: 412,816
: FILING DATE: 20-SEP-1989
: SEQ ID NO: 2
: LENGTH: 783
5422248-2

Query Match 15.8%; Score 57; DB 5; Length 783;
Best Local Similarity 24.0%; Pred. No. 24;
Matches 12; Conservative 9; Mismatches 23; Indels 6

QY 7 FQVDSRGAGCGQTSNDAAEAALNASSKEEAQITAAATPKYILFQFNL 56
DB 156 PTLKSFSGNCGTQGSILDCVPRQGSH-----CCIPRKHLLYQNM 199

RESULT 10
5422248-2
: Patent No. 5422248
: INVENTOR: CURTIS SMITH, CRAIG A.; LARSEN, ALF D.; SIMS, JOHN E.;
: BENSON, CURTIS
: TITLE OF INVENTION: DNA SEQUENCES ENCODING GRANULOCYTE-COLONY
: STIMULATING FACTOR RECEPTORS
: NUMBER OF SEQUENCES: 6
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/923,976
: FILING DATE: 15-JAN-1993
: PRIORITY CLAIM: YES
: APPLICATION NUMBER: 587,329
: FILING DATE: 24-SEP-1990
: APPLICATION NUMBER: 522,952
: FILING DATE: 03-APR-1990
: APPLICATION NUMBER: 416,306
: FILING DATE: 03-OCT-1989
: APPLICATION NUMBER: 412,816
: FILING DATE: 20-SEP-1989
: SEQ ID NO: 2
: LENGTH: 783
5422248-2

```



```

: TITLE OF INVENTION: Epilgrin, an Epithelial Ligand for
: OF SEQUENCES: 30
: NUMBER OF SEQUENCES: 30
: ADDRESSEE: Christensen, O'Connor, Johnson, and Kindness
: STREET: 1420 Fifth Avenue
: CITY: Seattle
: STATE: WA
: COUNTRY: USA
: ZIP: 98101-8100
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC Compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/600,982
: FILING DATE: 02-SEP-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Shelton, Dennis K.
: REGISTRATION NUMBER: 26,997
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 682-8100
: TELEFAX: (206) 224-0779
: INFORMATION FOR SEQ ID NO: 24:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1713 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: DESCRIPTION: E170 protein as translated from sequence
: of FIGURES 15A-15F, and as shown also in FIGURES
: US-08-600-982-24
:
: Query Match 15.1%; Score 54.5; DB 3; Length 1713;
: Best Local Similarity 32.8%; Pred. No. 1.5e+02;
: Matches 20; Conservative 11; Mismatches 23; Indels 7; Gaps 3;
:
: QY 9 VDCRGAGCGCCTSDNAAEALN--ASSKEALQ-IIAAIPKYLEFQPHLNSNLRIFD 65
: DB 553 VDAATAVENILMAKAEADNRAASASALQTVIKEDLPK----AKTSSNSDKLLN 608
:
: QY 66 K 66
: DB 609 E 609
:
: RESULT 14
: US-08-600-982-24
: Sequence 24, Application PC/TUS9410261A
: GENERAL INFORMATION:
: APPLICANT: Carter, William G.
: APPLICANT: Gil, Susanna A.
: APPLICANT: Ryan, Maureen C.
: TITLE OF INVENTION: Epilgrin, an Epithelial Ligand for
: OF SEQUENCES: 30
: NUMBER OF SEQUENCES: 30
: ADDRESSEE: Christensen, O'Connor, Johnson, and Kindness
: STREET: 1420 Fifth Avenue
: CITY: Seattle
: STATE: WA
: COUNTRY: USA
: ZIP: 98101-8100
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC Compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:

```

```

: APPLICATION NUMBER: PCT/US94/10261A
: FILING DATE: 02-SEP-1994
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Shelton, Dennis K.
: REGISTRATION NUMBER: 26,997
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 682-8100
: TELEFAX: (206) 224-0779
: INFORMATION FOR SEQ ID NO: 24:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1713 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: DESCRIPTION: E170 protein as translated from sequence
: of FIGURES 15A-15F, and as shown also in FIGURES 19A-19R
: PCT-US94-10261A-24
:
: Query Match 15.1%; Score 54.5; DB 4; Length 1713;
: Best Local Similarity 32.8%; Pred. No. 1.5e+02;
: Matches 20; Conservative 11; Mismatches 23; Indels 7; Gaps 3;
:
: QY 9 VDCRGAGCGCCTSDNAAEALN--ASSKEALQ-IIAAIPKYLEFQPHLNSNLRIFD 65
: DB 553 VDAATAVENILMAKAEADNRAASASALQTVIKEDLPK----AKTSSNSDKLLN 608
:
: QY 66 K 66
: DB 609 E 609
:
: RESULT 15
: US-08-276-919-10
: Sequence 24, Application US/08276919
: Patent No. 5585579
: GENERAL INFORMATION:
: APPLICANT: Torczynski, Richard M.
: APPLICANT: Bollon, Arthur P.
: TITLE OF INVENTION: Lung Cancer Marker
: NUMBER OF SEQUENCES: 16
: NUMBER OF SEQUENCES: 16
: ADDRESSEE: RICHARDS, MEDLOCK & ANDREWS
: STREET: 1201 Elm Street, Suite 4500
: CITY: Dallas
: STATE: TX
: COUNTRY: US
: ZIP: 75270-2197
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC Compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/276,919
: FILING DATE:
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Nolligan, Mark C.
: REGISTRATION NUMBER: 36,389
: REFERENCE/DOCKET NUMBER: B35792
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 214-939-4600
: TELEFAX: 214-939-4600
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 271 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-276-919-10

```

Query Match 15.0%; Score 54; DB 1; Length 271;  
Best local Similarity 29.5%; Pred. No. 15;  
Matches 13; Conservative 10; Mismatches 15; Indels 6; Gaps 1;  
OY 21 SNDAAEALNASSKEEALQITAAATPEKYLQPHHNLNSLDRIF 64  
DB 121 NSDLYPDASTASKSEGLAVI.V-----LIEMGSENPQYDKIF 158

Search completed: February 3, 2001, 02:19:43  
Job time: 170s sec

**THIS PAGE BLANK (USPTO)**



```

DR N-PSDB: T12904.
XX
XX Prodn. of virus-resistant transgenic plants - using mutated genomic
PT sequence from phytopathogenic DNA virus
XX
XX Disclosure: Fig 13; 93pp; French.
XX
CC Mutation of consensus amino acids in the NTP-binding site of
CC geminivirus Rep protein is used to produce replication deficient
CC viruses. The mutated viral nucleic acid is used for producing
CC transgenic plants that are resistant to, or tolerant of, the native
CC virus. The present sequence is a mutant form of the Rep (or C1)
CC virus, in which the wild-type Lys227 residue has been changed to a
CC His residue; transgenic Nicotiana benthamiana plants generated by
CC transformation with the mutated virus were not resistant to SPYLCV.
CC In contrast, plants transformed with a virus in which Lys227 had been
CC replaced by Ala were found to be resistant.
XX
XX Sequence 359 AA:
SQ
Query Match 61.5%; Score 222; DB 17; Length 359;
Best Local Similarity 58.0%; Pred. No. 5,1e-21;
Matches 40; Conservative 10; Mismatches 19; Indels 0; Gaps 0;
QY 2 LVNGEFQVDRSGCCTSDNAAEALNASSKEALQITAAAIPEKYLQFHNLSNLD 61
Db 111 LVNGTFLQGRSARGGQTANDAYAKALNAGSKQALVDIKELAPRDYVHFNHNSLD 170
QY 62 RIFDKTPEP 70
Db 171 KVFQVPPAP 179
RESULT 2
R88872
ID R88872 standard; Protein; 359 AA.
XX
XX R88871;
XX
XX 07-NOV-1996 (first entry)
XX
XX Sardinian tomato yellow leaf curl virus mutated Rep protein (K227H).
XX Nucleotide triphosphate binding site: DNA helicase; RNA helicase;
XX modification; mutation; viral replication; deficient; inhibition;
XX virus resistance; geminivirus; tomato yellow leaf curl virus;
XX Sardinian isolate; SPYLCV; transgenic plant; P-loop; C1 protein;
XX All protein.
XX
XX Sardinian tomato yellow leaf curl virus.
XX
XX Key Location/Qualifiers
XX Misc-difference 227
XX /note= "wild-type Lys has been replaced by His"
XX
XX WO9608573-A1.
XX
XX 21-MAR-1996.
XX
XX 15-SEP-1995; 95WO-FR01192.
XX
XX 15-SEP-1994; 94FR-0011040.
XX
XX (CNRS ) CENT NAT RECH SCI.
XX
XX Gronenborn B;
XX
XX WPI; 1996-179947/18.
XX
XX N-PSDB; T12905.
XX
XX Prodn. of virus-resistant transgenic plants - using mutated genomic
PT sequence from phytopathogenic DNA virus
XX
XX Disclosure: Fig 13; 93pp; French.

```

```

XX
XX Disclosure: Fig 13; 93pp; French.
XX
XX Mutation of consensus amino acids in the NTP-binding site of
XX geminivirus Rep protein is used to produce replication deficient
XX viruses. The mutated viral nucleic acid is used for producing
XX transgenic plants that are resistant to, or tolerant of, the native
XX virus. The present sequence is a mutant form of the Rep (or C1)
XX protein from the Sardinian isolate of tomato yellow leaf curl virus
XX (SPYLCV) in which the wild-type Lys227 residue has been changed to a
XX His residue; transgenic Nicotiana benthamiana plants generated by
XX transformation with the mutated virus were not resistant to SPYLCV.
XX In contrast, plants transformed with a virus in which Lys227 had been
XX replaced by Ala were found to be resistant.
XX
XX Sequence 359 AA:
SQ
Query Match 61.5%; Score 222; DB 17; Length 359;
Best Local Similarity 58.0%; Pred. No. 5,1e-21;
Matches 40; Conservative 10; Mismatches 19; Indels 0; Gaps 0;
QY 2 LVNGEFQVDRSGCCTSDNAAEALNASSKEALQITAAAIPEKYLQFHNLSNLD 61
Db 111 LVNGTFLQGRSARGGQTANDAYAKALNAGSKQALVDIKELAPRDYVHFNHNSLD 170
QY 62 RIFDKTPEP 70
Db 171 KVFQVPPAP 179
RESULT 3
R88872
ID R88872 standard; Protein; 359 AA.
XX
XX R88872;
XX
XX 07-NOV-1996 (first entry)
XX
XX Sardinian tomato yellow leaf curl virus mutated Rep protein (K227R).
XX Nucleotide triphosphate binding site: DNA helicase; RNA helicase;
XX modification; mutation; viral replication; deficient; inhibition;
XX virus resistance; geminivirus; tomato yellow leaf curl virus;
XX Sardinian isolate; SPYLCV; transgenic plant; P-loop; C1 protein;
XX All protein.
XX
XX Sardinian tomato yellow leaf curl virus.
XX
XX Key Location/Qualifiers
XX Misc-difference 227
XX /note= "wild-type Lys has been replaced by Arg"
XX
XX WO9608573-A1.
XX
XX 21-MAR-1996.
XX
XX 15-SEP-1995; 95WO-FR01192.
XX
XX 15-SEP-1994; 94FR-0011040.
XX
XX (CNRS ) CENT NAT RECH SCI.
XX
XX Gronenborn B;
XX
XX WPI; 1996-179947/18.
XX
XX N-PSDB; T12906.
XX
XX Prodn. of virus-resistant transgenic plants - using mutated genomic
PT sequence from phytopathogenic DNA virus
XX
XX Disclosure: Fig 13; 93pp; French.

```







```

RESULT 8
W34335
ID W34335 standard; Protein: 353 AA.
AC W34335;
XX
XX 27-APR-1998 (first entry)
XX Bean golden mosaic geminivirus CI BGA262 mutant.
XX
XX Geminivirus; BGWV; CI gene; transdominant mutation;
XX Transgenic plant; disease resistance.
XX
XX Bean golden mosaic virus type II isolate Guatemala.
XX
XX W09739110-A1.
XX
XX 23-OCT-1997.
XX
XX 15-APR-1997; 97WO-US06300.
XX
XX 16-APR-1996; 96US-0015517.
XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.
XX (WISC-) WISCONSIN ALUMNI RES FOUND.
XX
XX Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX
XX WPI: 1997-526447/48.
XX N-PSDB: T93293.
XX
XX Transgenic plants expressing geminivirus AC1 and CI wild-type and
XX mutant genes - have increased resistance to geminivirus infection
XX e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
XX golden mosaic geminivirus
XX
XX Example 5; Page 115-116; 132pp; English.
XX
XX This protein comprises a transdominant lethal mutant of the bean
XX golden mosaic virus (BGWV) CI protein (see W34338) that is required
XX for replication. It is encoded by mutated CI open reading frame
XX 606282 (see T93293), and carries a mutation in the NTP-binding
XX site. This mutant CI protein is transdominant lethal to transgenic
XX plants containing CI comprising geminivirus CI or AC1 wild-type or mutant
XX sequences that negatively interfere in trans with geminiviral
XX replication during infection. Such transgenic plants are resistant
XX to viral infection. The AC1/CI genes are especially from BGWV,
XX tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
XX and encode polypeptides (see W34324-35) that have mutations in the
XX highly conserved DNA-nicking and/or the NTP-binding domains.
XX
XX Sequence 353 AA;

Query Match 60.7%; Score 219; DB 18; Length 353;
Best Local Similarity 60.0%; Pred. No. 1.2e-20;
Matches 42; Conservative 11; Mismatches 17; Indels 0; Gaps 0;

Oy 1 TLVWGFQVQGRSGCGCTSDNAAALNNSKEEALQIIAAALPEKYLFOFHNLNSNL 60
Db 110 TLVGFQVQGRSGCGCTSDNAAALNNSKEEALQIIAAALPEKYLFOFHNLNSNL 169
Oy 61 DRIFDKTPEP 70
Db 170 erifvkepep 179

RESULT 9
P70407
ID P70407 standard; Protein: 353 AA.
XX
XX Query Match 60.1%; Score 217; DB 6; Length 353;
XX Best Local Similarity 60.0%; Pred. No. 1.2e-20;
XX Matches 42; Conservative 11; Mismatches 17; Indels 0; Gaps 0;

Oy 1 TLVWGFQVQGRSGCGCTSDNAAALNNSKEEALQIIAAALPEKYLFOFHNLNSNL 60
Db 110 TLVGFQVQGRSGCGCTSDNAAALNNSKEEALQIIAAALPEKYLFOFHNLNSNL 169
Oy 61 DRIFDKTPEP 70
Db 170 erifvkepep 179

RESULT 10
W34336
ID W34336 standard; Protein: 361 AA.
XX
XX W34336;
XX
XX 27-APR-1998 (first entry)
XX
XX Tomato mottle virus AC1 protein.
XX
XX Geminivirus; TOMOV; AC1 gene; transdominant mutation;
XX Transgenic plant; disease resistance.
XX
XX Tomato mottle virus isolate Florida.
XX
XX W09739110-A1.
XX
XX 23-OCT-1997.
XX
XX 15-APR-1997; 97WO-US06300.
XX
XX 16-APR-1996; 96US-0015517.
XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.
XX (WISC-) WISCONSIN ALUMNI RES FOUND.
XX

```



PI transgenic plants expressing geminivirus AC1 and CI WTD-type and

CC. The sequence  
CC strand of the

CC. The sequence is encoded by ORF 4 which occurs on the complementary  
CC strand of the (b) molecule of the geminivirus.  
CC See also p70559-p70567.

XX SQ Sequence 361 AA;

Query Watch 56.4%; Score 203.5; DB 8; Length 361;  
 Best Local Similarity 63.1%; Pred. No. 1.4e-18;  
 Matches 41; Conservative 9; Mismatches 14; Indels 1; Gaps 1;  
 XX 5 GEFQVDSRGSGCSTNDAAAEALNASSKEALQITAAAIPEKYLFOFHNLSNDRIF 64  
 Db 114 GSFQVDSRGSGCSTNDAAAEALNASSKEALQITAAAIPEKYLFOFHNLSNDRIF 173  
 QY 65 -DKTP 68  
 Db 174 adevp 178

# RESULT 15

XX W34329 standard; Protein; 357 AA.

XX W34329;

DT 27-APR-1998 (first entry)

DE Tomato yellow leaf curl virus mutant CI protein pTYK104R #1.

KW Geminivirus; pTYK104R #1; CI protein; transdominant mutation;

XX transgenic plant; disease resistance.

OS Tomato yellow leaf curl virus strain Israel.

XX WO9739110-A1.

XX WO9739110-A1.

PD 23-OCT-1997.

PF 15-APR-1997; 97WO-0506300.

PR 16-APR-1996; 96US-0015517.

PA (SEMI-) SEMINIS VEGETABLE SEEDS INC.

XX (WISC) WISCONSIN ALUMNI RES FOUND.

PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;

DR WPI; 1997-526447/48.

DR N-PSDB; T93287.

XX Transgenic plants expressing geminivirus AC1 and CI wild-type and

PT mutant genes -- have increased resistance to geminivirus infection

PT e.g. tomato mottle virus tomato yellow leaf curl virus or bean

XX golden mosaic geminivirus

PS Example 4.5; Page 84-86; 132pp; English.

XX This protein comprises a transdominant lethal mutant of the tomato

CC (Tomato leaf curl virus (TLCV) CI protein (see W34337) that is

CC required for replication of the virus. It is encoded by the CI gene

CC (see T93287) of an infectious clone of TYLCV and carries a single

CC mutation in the DNA-nicking domain. The invention involves

CC production of transgenic plants containing DNA comprising

CC geminivirus CI or AC1 wild-type or mutant sequences that negatively

CC interfere in trans with geminiviral replication during infection.

CC Such transgenic plants are resistant to viral infection. The

CC W34329 genes are especially from tomato mottle virus or

CC bean yellow mosaic virus (see T93287) and have been encoded

CC polypeptides (see W34324-35) that have mutations in the highly

CC conserved DNA-nicking and/or NTP-binding domains.

XX Sequence 357 AA;

Query Match 55.4%; Score 200; DB 18; Length 357;  
 Best Local Similarity 63.9%; Pred. No. 3.9e-18;  
 Matches 39; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

QY 4 WGEFQVDSRGSGCSTNDAAAEALNASSKEALQITAAAIPEKYLFOFHNLSNDRIF 63  
 Db 111 GSFQVDSRGSGCSTNDAAAEALNASSKEALQITAAAIPEKYLFOFHNLSNDRIF 170

QY 64 F 64

Db 171 f 171

Search completed: February 3, 2001, 02:15:27  
 Job time: 3434 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: February 3, 2001, 02:15:27 ; Search time 144.12 seconds  
(without alignments)  
16.608 Million cell updates/sec

Title: US-09-289-346a-5  
Perfect score: 359  
Sequence: 1 TLVNGEFVDGRSARGQQT.....TQFHINSLDRIFDKTEEP 70

Scoring table:  
BLOSUM62  
Gap 10.0, Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A.Geneseq\_36:\*

- 1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT:\*
- 2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT:\*
- 3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT:\*
- 4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT:\*
- 5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT:\*
- 6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT:\*
- 7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT:\*
- 8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT:\*
- 9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT:\*
- 10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT:\*
- 11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT:\*
- 12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT:\*
- 13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT:\*
- 14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT:\*
- 15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT:\*
- 16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT:\*
- 17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT:\*
- 18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT:\*
- 19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT:\*
- 20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT:\*
- 21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT:\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	220	61.3	361	18 W34336	Tomato mottle virus
2	220	61.3	361	18 W34324	Tomato mottle virus
3	220	61.3	361	18 W34325	Tomato mottle virus
4	220	61.3	361	18 W34326	Tomato mottle virus
5	219	61.0	359	17 R88870	Sardinian tomato Y
6	219	61.0	359	17 R88871	Sardinian tomato Y
7	219	61.0	359	17 R88872	Sardinian tomato Y
8	218	60.7	353	18 W34338	Bean golden mosaic
9	218	60.7	353	18 W34332	Bean golden mosaic
10	218	60.7	353	18 W34333	Bean golden mosaic
11	218	60.7	353	18 W34334	Bean golden mosaic
12	218	60.7	353	18 W34335	Bean golden mosaic

ORF 4 gene product  
Product of ORF 4 f  
Tobacco leaf curl  
Tomato yellow leaf  
Tomato yellow leaf  
Tomato yellow leaf  
Tomato yellow leaf  
HIV-1 strain YBF30  
Deduced sequence o  
Tomato mottle viru  
Tomato mottle viru  
Membrane bound pro  
Human secreted pro  
Human secreted pro  
Human calcium chan  
Human calcium chan  
Sequence of the al  
Human neuronal cal  
Human neuronal cal  
Human calcium chan  
Human calcium chan  
Human calcium chan  
Human calcium chan  
Mutant Aspergillus  
A. oryzae PA-8.1 l  
Virulence gene clu  
Human 5' EST seque  
Geminivirus Rep C1  
Mastrevirus Rep C1  
Lettuce resistance  
Lettuce resistance  
Amino acid sequenc  
Prepro-polygalactu  
Solanum melongena  
Amino acid sequenc

## ALIGNMENTS

RESULT 1  
W34336 ID W34336 standard; Protein; 361 AA.  
XX AC W34336;  
XX AC W34336;  
XX DT 27-APR-1998 (first entry)  
XX DE Tomato mottle virus AC1 protein.  
XX GM Geminivirus; ToMoV; AC1 gene; transdominant mutation;  
XX KW Transgenic plant; disease resistance.  
XX XX Tomato mottle virus isolate Florida.  
XX XX W05739110-A1.  
XX XX 23-OCT-1997.  
XX XX 15-APR-1997; 97MO-US06300.  
XX XX 16-APR-1997; 96US-0015517.  
XX XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.  
XX XX (WISC) WISCONSIN ALUMNI RES FOUND.  
XX XX Ahlquist PG, Hanson SF, Lau HT, Maxwell DP, Stout JT;  
XX WPI: 1997-52647/48.  
XX N-PSDB; 193294.

Transgenic plants expressing geminivirus AC1 and C1 wild-type and  
mutant genes - have increased resistance to geminivirus infection  
e.g. tomato mottle virus, tomato yellow leaf curl virus or bean  
golden mosaic geminivirus

**Example 3.3:** Page 57-58; 132pp; English.

This protein comprises the wild-type AC1 protein of tomato mottle virus (ToMoV), a geminivirus that has a bipartite genome. The AC1 gene (see T93294) must be expressed for efficient replication of the two genomic components, DNA-A and DNA-B, of the bipartite ToMoV genome. The invention involves production of transgenic plants containing DNA comprising the AC1 gene of the wild-type ToMoV. Such transgenic plants interfere in trans with geminiviral replication during infection. Such transgenic plants negatively interfere in trans or wild-type or mutant sequences that contain AC1 or C1 with geminiviral replication during infection. Such transgenic plants are resistant to viral infection. The AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl virus or bean golden mosaic geminivirus (see T93282-93) and encode polypeptides (see W34324-35) that have mutations in the highly conserved DNA-nicking and/or NTP-binding domains.

Query Match 61.3%; Score 220; DB 18; Length 361;  
Best Local Similarity 58.6%; Pred. No. 1.4e-20;  
Matches 41; Conservative 14; Mismatches 15; Indels 0; Gaps 0;

OY	1	TLMWGEPQVDGRAGGCGTSDNDAAEALNAKSEALOTIREKIPAAALPFGHNLNML 60
DB	110	tiewgdfqldgragggcsandysakalnassvgsalavireeqpkdflqphnrsnl 169
OY	61	DRIFFKTPEP 70
DB	170	erifakapep 179
OY	61	
DB	170	

RESULT 3  
W34325

ID	W34325 standard; Protein: 361 AA.
AC	W34325;
XX	
XX	27-APR-1998 (first entry)
DT	
DE	Tomato mottle virus AC1 mutant TOMV-AC1dlm1.
DE	
KW	Geminivirus: ToMoV-AC1dlm1; AC1 gene: transdominant mutation;
KW	transgenic plant; disease resistance.
OS	Tomato mottle virus isolate Florida.
OS	Synthetic.
PN	W09739110-Al.
PN	
PD	23-OCT-1997.
XX	
XX	15-APR-1997; 97WO-USO6300.
PF	
XX	
XX	16-APR-1996; 96US-0015517.
XX	
PA	(SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA	(WISC ) WISCONSIN ALUMNI RES FOUND.
PI	Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX	
DR	WPI: 1997-526447/48.
NF	N-PDSB; T93283.
XX	
PT	Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT	mutant genes - have increased resistance to geminivirus infection
PT	e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT	golden mosaic geminivirus
PT	
PS	Example 3.5; Page 64-65; 132pp; English.
CC	This protein comprises a transdominant lethal mutant, designated
CC	TOMOV-AC1dlm1, of tomato mottle virus (ToMoV) AC1 protein (see

**Example 3.4:** Page 60-62; 132pp; English.

This protein comprises the wild-type AC1 protein of tomato mottle virus (ToMoV), a geminivirus that has a bipartite genome. The AC1 gene (see T93294) must be expressed for efficient replication of the two genomic components, DNA-A and DNA-B. The AC1 protein has a high level of sequence identity with the AC1 protein of other geminiviruses, and an NTP binding activity. The invention involves production of transgenic plants containing DNA comprising AC1 or C1 wild-type or mutant sequences that negatively interfere in trans with geminiviral replication during infection. Such transgenic plants are resistant to viral infection. The AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl virus or bean golden mosaic geminivirus (see T93282-93) and encode polypeptides (see W34324-35) that have mutations in the highly conserved DNA-nicking domain and/or the NTP-binding domain.

Query Match 61.3%; Score 220; DB 18; Length 361;  
Best Local Similarity 58.6%; Pred. No. 1.4e-20;  
Matches 41; Conservative 14; Mismatches 15; Indels 0; Gaps 0;

OY	1	TLMWGEPQVDGRAGGCGTSDNDAAEALNAKSEALOTIREKIPAAALPFGHNLNML 60
DB	110	tiewgdfqldgragggcsandysakalnassvgsalavireeqpkdflqphnrsnl 169
OY	61	DRIFFKTPEP 70
DB	170	erifakapep 179
OY	61	
DB	170	

RESULT 2  
W34324

ID	W34324 standard; Protein: 361 AA.
AC	W34324;
XX	
XX	27-APR-1998 (first entry)
DT	
DE	Tomato mottle virus AC1 mutant TOMV-AC1dlm1.
DE	
KW	Geminivirus: ToMoV-AC1dlm1; AC1 gene: transdominant mutation;
KW	transgenic plant; disease resistance.
OS	Tomato mottle virus isolate Florida.
OS	Synthetic.
PN	W09739110-Al.
PN	
PD	23-OCT-1997.
XX	
XX	15-APR-1997; 97WO-USO6300.
PF	
XX	
XX	16-APR-1996; 96US-0015517.
PR	
XX	(SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA	(WISC ) WISCONSIN ALUMNI RES FOUND.
PA	
PI	Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX	
DR	WPI: 1997-526447/48.
NF	N-PDSB; T93282.
XX	
PT	Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT	mutant genes - have increased resistance to geminivirus infection
PT	e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT	golden mosaic geminivirus
PT	
PS	Example 3.4; Page 60-62; 132pp; English.
XX	



CC CC It is encoded by a mutant AC1 gene (see P93283) of  
CC CC tomato mottle virus (Tomev), and carries a mutation in an  
CC CC NTP-binding domain. The AC1 gene (see also P93294) must be  
CC CC expressed for efficient replication of the two genomic components,  
CC CC DNA-A and DNA-B, of the bipartite Tomato genome. The invention  
CC CC involves production of transgenic plants containing DNA comprising  
CC CC geminivirus AC1 or CI wild-type or mutant sequences that negatively  
CC CC interfere in trans with geminiviral replication during infection.  
CC CC Such transgenic plants are resistant to viral infection.  
CC CC Such transgenic plants are especially resistant to tomato leaf curl  
CC CC virus or to any other virus (see P93283).  
CC CC polypeptides (see P4324-35) that have mutations in the highly  
CC CC conserved DNA-nicking and/or NTP-binding domains.  
XX Sequence 361 AA;

Query Match 61.3%; Score 220; DB 18; Length 361;  
Best Local Similarity 58.6%; Pred. No. 1.4e-20;  
Matches 41; Conservative 14; Mismatches 15; Indels 0; Gaps

Oy	1	TLWNCPEVDGKSRGCGTOSNDAAEALNASSKEALOTREKIDPAALFOFHLSNL	60
DB	110	tlwncdqlqldrgsrqgsandysakalnassvsgalavireeqpkfvglnmrsl	169
Oy	61	DRFDTPPEP	70
DB	170	erfkapep	179

	RESULT	4
	W34326	
	ID	W34326 standard; Protein; 361 AA.
	XX	
	AC	
	W34326;	
	XX	
	DT	27-APR-1998 (first entry)

KW	Geminivirus; ToMoV-AcidIm23; AC1 gene; transdominant mutation;
KW	transgenic plant; disease resistance.
KW	
OS	Tomato mottle virus isolate Florida.
OS	Synthetic.
XX	
PN	WO9739110-A1.

PD	23-OCT-1997.	
XX		
PF	15-APR-1997;	97MO-US06300.
XX		
PR	16-APR-1996;	96US-0015517.
XX		
XX	(SEMI-) SEMINIS VEGETABLE SEEDS INC.	
PA	(WISC.) WISCONSIN ALUMNT. RES. FOUND.	

XX Abliquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;  
PI  
XX  
DR WPI: 1997-526447/48.  
DR N-PSDB: T93384.END.

Transgenic plants expressing geminivirus AC1 and CI wild-type and mutant genes - have increased resistance to geminivirus infection e.g. tomato mottle virus, tomato yellow leaf curl virus or bean golden mosaic geminivirus

PS Example 3.6; Page 67-69; 132pp; English.

This protein comprises a transdominant lethal mutant, designated TOMOV-AC1dln23, of tomato mottle virus (TOMOV) AC1 protein (see W34336). It is encoded by a mutant AC1 gene (see T93284) of CC tomato mottle virus (TOMOV), and carries 2 mutations in an CC

NTP-binding domain. The ACl gene (see also Tg3294), must be expressed for efficient replication of the two genomic components. DNA-A and DNA-B, of the bipartite tomato genome. The invention involves production of transgenic plants containing DNA comprising geminivirus ACl or CI wild-type or mutant sequences that negatively interfere in trans with geminiviral replication during infection. Such transgenic plants are resistant to viral infection. The ACl/CI genes are especially from Tomato, tomato yellow leaf curl virus or bean golden mosaic geminivirus (see Tg3292-93) and encode highly conserved domains. The invention provides mutations in the highly conserved DNA-A/RCKing and/or NTP-binding domains.

Sequence 361 AA;  
XX

Query Match 61.3%; Score 220; DB 18; Length 361;  
Best Local Similarity 58.6%; Pred. No. 1.4e-20;  
Matches 41; Conservative 14; Mismatches 15; Indels 0; Gaps

[illegible]

RESULT 5  
R88870  
ID R88870 standard: Protein: 359 AA.

AC R88870;  
XX  
DT 07-NOV-1996 (first entry)

DE Sardinian tomato yellow leaf curl virus mutated Rep protein (K227A).

Nucleotide triphosphate binding site: DNA helicase;  
 KW modification; mutation: viral replication; deficient: inhibition;  
 KW viral resistance; geminivirus: tomato yellow leaf curl virus;  
 KW Sardinian isolate: STYLCO; transgenic plant; P-loop; CI protein;  
 KW ALI protein; dominant negative phenotype.  
 XX  
 OS Sardinian tomato yellow leaf curl virus.

EH	Key	Location/Qualifiers
FT	Misc-difference	227
FT	FT	/note= "wild-type Lys has been replaced by Ala"
XX		
PN		W09608573-Al.

PD 21-MAR-1996.

PF 15-SEP-1995; 95WO-FR01192.

PR 15-SEP-1994; 94FR-0011040.

PA (CNRS ) CENT NAT RECH SCI.

PI Gronenborn B;

DR WPI; 1996-179947/18.

[illegible]

PT  
sequence from phytopathogenic DNA virus  
XX

DISCLOSURE; FIG 13; 33PP; FRENCH.

CC mutation of consensus amino acids in the NTP-binding site of  
CC  
CC geminivirus Rep protein is used to produce replication deficient

[illegible]



CC replaced by Ala were found to be resistant.

XX Sequence 359 AA;

Query Match 61.0%; Score 219; DB 17; Length 359;

Best Local Similarity 58.0%; Pred. No. 1.8e-20;

Matches 40; Conservative 11; Mismatches 18; Indels 0; Gaps 0;

Qy 2 TLWGEFQVDSRGCGCQTSNDAAAEALNASSKEALQIREKIPAAALFOFHNLNSNLD 61

Db 111 TLWGEFQVDSRGCGCQTSNDAAAEALNASSKEALQIREKIPAAALFOFHNLNSNLD 170

Qy 62 DRFDKTPPEP 70

Db 171 kvfvpvpp 179

RESULT 8

W34338

ID W34338 standard; Protein; 353 AA.

XX

AC W34338;

XX 27-APR-1998 (first entry)

XX Bean golden mosaic geminivirus CI protein.

XX Geminivirus; BGWV; CI gene; transdominant mutation;

XX Transgenic plant; disease resistance.

XX Bean golden mosaic virus type II isolate Guatemala.

XX W09739110-AL.

XX 23-OCT-1997.

XX 15-APR-1997; 97WO-US06300.

XX 16-APR-1996; 96US-0015517.

XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.

XX (WISC ) WISCONSIN ALJMNIT RES FOUND.

XX Abhiquist PG, Hanson SF, Lau HT, Maxwell DP, Stout JT;

XX WPI: 1997-526447/48.

XX N-PSDB: T93314.

XX Transgenic plants expressing geminivirus AC1 and CI wild-type and

XX mutant genes - have increased resistance to geminivirus infection

XX e.g. tomato mottle virus, tomato yellow leaf curl virus or bean

XX golden mosaic geminivirus

XX Example 5; Page 100-102; 132pp; English.

XX This sequence comprises the bean golden mosaic virus (BGWV) CI

XX protein that is required for replication. The invention involves

XX a method of transgenic plants containing RNA comprising CI or AC1

XX wild-type or mutant genes that are resistant to infection by such

XX wild-type or mutant genes during infection. Such transgenic

XX plants are resistant to viral infection. The AC1/CI genes are

XX especially from BGWV, tomato mottle virus or tomato yellow leaf

XX curl virus (see T93282-93) and encode polypeptides (see W34324-35)

XX that have mutations in the highly conserved DNA-nicking domain

XX and/or the NTP-binding domains.

XX Sequence 353 AA;

Query Match 60.7%; Score 218; DB 18; Length 353;

Best Local Similarity 60.0%; Pred. No. 2.4e-20;

Matches 42; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

Qy 1 TLWGEFQVDSRGCGCQTSNDAAAEALNASSKEALQIREKIPAAALFOFHNLNSNLD 60

Db 110 TLWGEFQVDSRGCGCQTSNDAAAEALNASSKEALQIREKIPAAALFOFHNLNSNLD 169

Qy 61 DRFDKTPPEP 70

Db 170 erifvkvpep 179

RESULT 9

W34332

ID W34332 standard; Protein; 353 AA.

XX

AC W34332;

XX 27-APR-1998 (first entry)

XX Bean golden mosaic geminivirus CI BGAL90 mutant.

XX Geminivirus; BGWV; CI gene; transdominant mutation;

XX Transgenic plant; disease resistance.

XX Bean golden mosaic virus type II isolate Guatemala.

XX W09739110-AL.

XX 23-OCT-1997.

XX 15-APR-1997; 97WO-US06300.

XX 16-APR-1996; 96US-0015517.

XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.

XX (WISC ) WISCONSIN ALJMNIT RES FOUND.

XX Abhiquist PG, Hanson SF, Lau HT, Maxwell DP, Stout JT;

XX WPI: 1997-526447/48.

XX N-PSDB: T93290.

XX Transgenic plants expressing geminivirus AC1 and CI wild-type and

XX mutant genes - have increased resistance to geminivirus infection

XX e.g. tomato mottle virus, tomato yellow leaf curl virus or bean

XX golden mosaic geminivirus

XX Example 5; Page 103-105; 132pp; English.

XX This protein comprises a control mutant of the bean golden mosaic

XX virus (BGWV) CI protein (see W34338) that is required for

XX replication. It is encoded by mutated CI open reading frame

XX (see T93290). The invention involves production of or AC1

XX wild-type or mutant genes that are resistant to infection by such

XX wild-type or mutant genes during infection. Such transgenic

XX plants are resistant to viral infection. The AC1/CI genes are

XX especially from BGWV, tomato mottle virus or tomato yellow leaf

XX curl virus (see T93282-93) and encode polypeptides (see W34324-35)

XX that have mutations in the highly conserved DNA-nicking and/or the

XX NTP-binding domains.

XX Sequence 353 AA;

Query Match 60.7%; Score 218; DB 18; Length 353;

Best Local Similarity 60.0%; Pred. No. 2.4e-20;

Matches 42; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

Qy 1 TLWGEFQVDSRGCGCQTSNDAAAEALNASSKEALQIREKIPAAALFOFHNLNSNLD 60

Db 110 TLWGEFQVDSRGCGCQTSNDAAAEALNASSKEALQIREKIPAAALFOFHNLNSNLD 169

Qy 61 DRFDKTPPEP 70

Db 170 erifvkvpep 179

```

Db 170 erifvkpvp 179
      :||| | |||
RESULT 10
W34333
ID W34333 standard; Protein: 353 AA.
AC W34333;
XX 27-APR-1998 (first entry)
XX Bean golden mosaic geminivirus C1 BGA221 mutant.
DE Geminivirus; BGWV; C1 gene; transdominant mutation;
KW transgenic plant; disease resistance.
XX Bean golden mosaic virus type II isolate Guatemala.
OS WO9739110-A1.
PN WO9739110-A1.
PD 23-OCT-1997.
PE 15-APR-1997; 97WO-US06300.
PR 16-APR-1996; 96US-0015517.
XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA (WISC ) WISCONSIN ALUMNI RES FOUND.
XX Ahlquist FG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
DR WPI: 1997-526447/48.
DR N-PSDB: T93291.
XX Transgenic plants expressing geminivirus AC1 and CI wild-type and
PT mutant genes - have increased resistance to geminivirus infection
PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT golden mosaic geminivirus
XX Example 5; Page 107-109; 132pp; English.
CC This protein comprises a transdominant lethal mutant of the bean
CC golden mosaic virus (BGWV) C1 protein (see W34338) that is required
CC for replication. It is encoded by mutated C1 open reading frame
CC BGA221 (see T93291) and carries a mutation in the NTP-binding
CC domain. The invention involves production of transgenic plants
CC containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC sequences that negatively interfere in trans with geminivirus
CC replication and/or the NTP-binding domain. Such transgenic plants are resistant
CC to viral infection. The AC1/C1 genes are especially from BGWV
CC tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC and encode polypeptides (see W34324-35) that have mutations in the
CC highly conserved DNA-nicking and/or the NTP-binding domains..
XX Sequence 353 AA;
S Query Match 50.7%; Score 218; DB 18; Length 353;
Best Local Similarity 60.0%; Pred. No. 2.4e-20;
Matches 42; Conservative 12; Mismatches 16; Indels 0; Gaps 0;
QY 1 TLVWGFQVDRSGRGCGTSDNDAAAHALNASSKEALQIREKIPAAALFQFINLSNL 60
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 110 TLWGGTQVDRSGRGCGTSDNDAAAHALNASSKEALQIREKIPAAALFQFINLSNL 169
      :||| | |||
QY 61 DRIFDKTPEP 70
      :||| | |||
Db 170 erifvkpvp 179
      :||| | |||
RESULT 12
W34335
ID W34335 standard; Protein: 353 AA.
AC W34335;
XX 27-APR-1998 (first entry)
XX
XX
XX

```

```

ID W34334 standard; Protein: 353 AA.
XX
XX W34334;
XX 27-APR-1998 (first entry)
XX Bean golden mosaic geminivirus C1 BGA228 mutant.
DE Geminivirus; BGWV; C1 gene; transdominant mutation;
KW transgenic plant; disease resistance.
XX Bean golden mosaic virus type II isolate Guatemala.
OS WO9739110-A1.
PN WO9739110-A1.
PD 23-OCT-1997.
PE 15-APR-1997; 97WO-US06300.
PR 16-APR-1996; 96US-0015517.
XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA (WISC ) WISCONSIN ALUMNI RES FOUND.
XX Ahlquist FG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
DR WPI: 1997-526447/48.
DR N-PSDB: T93292.
XX Transgenic plants expressing geminivirus AC1 and CI wild-type and
PT mutant genes - have increased resistance to geminivirus infection
PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT golden mosaic geminivirus
XX Example 5; Page 111-112; 132pp; English.
CC This protein comprises a transdominant lethal mutant of the bean
CC golden mosaic virus (BGWV) C1 protein (see W34338) that is required
CC for replication. It is encoded by mutated C1 open reading frame
CC BGA228 (see T93292) and carries a mutation in the NTP-binding
CC domain. The invention involves production of transgenic plants
CC containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC sequences that negatively interfere in trans with geminivirus
CC replication during infection. Such transgenic plants are resistant
CC to viral infection. The AC1/C1 genes are especially from BGWV
CC tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC and encode polypeptides (see W34324-35) that have mutations in the
CC highly conserved DNA-nicking and/or the NTP-binding domains..
XX Sequence 353 AA;
S Query Match 60.7%; Score 218; DB 18; Length 353;
Best Local Similarity 60.0%; Pred. No. 2.4e-20;
Matches 42; Conservative 12; Mismatches 16; Indels 0; Gaps 0;
QY 1 TLVWGFQVDRSGRGCGTSDNDAAAHALNASSKEALQIREKIPAAALFQFINLSNL 60
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 110 TLWGGTQVDRSGRGCGTSDNDAAAHALNASSKEALQIREKIPAAALFQFINLSNL 169
      :||| | |||
QY 61 DRIFDKTPEP 70
      :||| | |||
Db 170 erifvkpvp 179
      :||| | |||
RESULT 12
W34335
ID W34335 standard; Protein: 353 AA.
AC W34335;
XX 27-APR-1998 (first entry)
XX
XX
XX

```

DE Bean golden mosaic geminivirus CI BGA262 mutant.  
 XX Geminivirus; BGWV; CI gene; transdominant mutation;  
 KW Transgenic plant; disease resistance.  
 XX  
 OS Bean golden mosaic virus type II isolate Guatemala.  
 XX  
 FN W09739110-A1.  
 XX  
 PD 23-OCT-1997.  
 XX  
 XX 15-APR-1997; 97NO-US06300.  
 XX  
 PR 16-APR-1996; 96US-0015517.  
 XX  
 XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.  
 FA (WISC) WISCONSIN AUMNI RES FOUND.  
 XX  
 XX Ahlquist PG, Hanson SF, Lau HT, Maxwell DP, Stout JT:  
 XX N-PSDB: T93293.  
 XX  
 XX Transgenic plants expressing geminivirus AC1 and C1 wild-type and  
 PT mutant genes - have increased resistance to geminivirus infection  
 PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean  
 PT golden mosaic geminivirus  
 XX  
 PS Example 5; Page 115-116; 132pp; English.  
 XX  
 CC This protein comprises a transdominant lethal mutant of the bean  
 CC golden mosaic virus (BGWV) C1 protein (see W34338) that is required  
 CC for replication. It is encoded by mutated C1 open reading frame  
 CC BGAC252 (see T93293) and carries a mutation in the NTP-binding  
 CC containing DNA nicks. Involves production of transgenic plants  
 CC sequences that negatively interfere in trans with geminiviral  
 CC replication during infection. Such transgenic plants are resistant  
 CC to viral infection. The AC1/C1 genes are especially from BGWV,  
 CC tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)  
 CC and encode polypeptides (see W34324-35) that have mutations in the  
 CC highly conserved DNA-nicking and/or the NTP-binding domains.  
 XX  
 SQ Sequence 353 AA:

Query Match 60.7%; Score 218; DB 18; Length 353;  
 Best Local Similarity 60.0%; Pred. No. 2.4e-20;  
 Matches 42; Conservative 12; Mismatches 16; Indels 0;

Oy 1 TLVNGEPOVGRARGCOTNDAAAEALNASKEEALQIIREKTPAAALFOPHNLNSLI 60  
 :|||||  
 Db 110 tiewggfvdgrarggggsandsyakaalnadsiesaitlkeepkdyvighnairsnl 169  
 :|||||

Oy 61 DRIFDKTPEP 70  
 :|||||  
 Db 170 eritfvkvppep 179  
 :|||||

RESULT 13  
 P70407  
 ID P70407 standard; Protein: 353 AA.  
 AC  
 AC P70407:  
 DE 30-APR-1991 (first entry)  
 XX  
 XX Product of ORF 4 from Bean Golden Mosaic virus.  
 XX  
 XX Mungbean yellow mosaic virus.  
 OS  
 OS Mungbean yellow mosaic virus.  
 FN JP62126982-A.  
 XX  
 PD 09-JUN-1987.  
 XX  
 XX 28-NOV-1985; 85JP-0266080.  
 XX  
 XX 28-NOV-1985; 85JP-0266080.  
 XX  
 XX (TEIJ ) TEIJIN KK.  
 XX  
 XX WPI: 1987-196308/28.  
 DR N-PSDB: N70895.  
 XX  
 XX Novel DNA and hybrid DNA useful - as vector for recombinant work  
 PT of plant gene.  
 XX  
 XX Disclosure: Fig 8. 21pp; Japanese.  
 XX  
 CC The sequence is encoded by ORF 4 which occurs on the complementary  
 CC strand of the (b) molecule of the geminivirus.  
 CC See also P70559-P70567.

PN JP61257186-A.  
 XX  
 PD 14-NOV-1986.  
 XX  
 XX 10-MAY-1985; 85JP-0098108.  
 XX  
 PR 10-MAY-1985; 85JP-0098108.  
 XX  
 XX (TEIJ ) TEIJIN KK.  
 XX  
 XX WPI: 1987-159662/23.  
 DR N-PSDB: N70630.  
 XX  
 XX New DNA and hybrid DNA - used for recombinant vector of plants.  
 PT  
 XX  
 XX Disclosure: Fig 6; 24pp; Japanese.  
 PS  
 XX The sequence encoding this protein may be taken along with the -a  
 CC DNA sequence and a suitable resistance gene, and used to create a  
 CC recombinant plant vector.  
 CC See also N70623.  
 XX  
 SQ Sequence 353 AA:

Query Match 60.2%; Score 216; DB 8; Length 353;  
 Best Local Similarity 60.0%; Pred. No. 4.4e-20;  
 Matches 42; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

Oy 1 TLVNGEPOVGRARGCOTNDAAAEALNASKEEALQIIREKTPAAALFOPHNLNSLI 60  
 :|||||  
 Db 110 tiewggfvdgrarggggsandsyakaalnadsiesaitlkeepkdyvighnairsnl 169  
 :|||||

Oy 61 DRIFDKTPEP 70  
 :|||||  
 Db 170 eritfvkvppep 179  
 :|||||

RESULT 14  
 P70562  
 ID P70562 standard; Protein: 361 AA.  
 AC  
 AC P70562:  
 DE 30-APR-1991 (first entry)  
 XX  
 XX Product of ORF 4 from MYMV complementary strand (b).  
 XX  
 XX Geminivirus.  
 XX  
 OS Mungbean yellow mosaic virus.  
 FN JP62126982-A.  
 XX  
 PD 09-JUN-1987.  
 XX  
 XX 28-NOV-1985; 85JP-0266080.  
 XX  
 XX 28-NOV-1985; 85JP-0266080.  
 XX  
 XX (TEIJ ) TEIJIN KK.  
 XX  
 XX WPI: 1987-196308/28.  
 DR N-PSDB: N70895.  
 XX  
 XX Novel DNA and hybrid DNA useful - as vector for recombinant work  
 PT of plant gene.  
 XX  
 XX Disclosure: Fig 8. 21pp; Japanese.  
 XX  
 CC The sequence is encoded by ORF 4 which occurs on the complementary  
 CC strand of the (b) molecule of the geminivirus.  
 CC See also P70559-P70567.

Search completed: February 3, 2001, 02:15:28  
Job time: 3435 sec

GenCode version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: February 3, 2001, 02:19:43 : Search time 109.09 Seconds  
(without alignments)  
11.523 Million cell updates/sec

Title: US-09-289-346a-5  
Perfect score: 359  
Sequence: 1 TLVNGEFQVDSARGCOT.....QFHNLSNLDRIPTKPE 70

Scoring table: BLOSUM62  
Gap 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cqn2\_6/ptodata/1/aa/5A\_COMB.pep.\*  
2: /cqn2\_6/ptodata/1/aa/5B\_COMB.pep.\*  
3: /cqn2\_6/ptodata/1/aa/6\_COMB.pep.\*  
4: /cqn2\_6/ptodata/1/aa/PCTUS\_COMB.pep.\*  
5: /cqn2\_6/ptodata/1/aa/backfilled.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result N	Score	Query Match %	Length	DB ID	Description
1	219	61.0	359	3	US-08-609-103B-2
2	219	61.0	359	3	US-08-609-103B-4
3	219	61.0	359	3	US-08-609-103B-6
4	219	61.0	359	3	US-08-609-103B-8
5	63.5	17.7	1713	3	US-08-609-382-24
6	53.2	16.6	2163	4	PCT-US95-026114-24
7	53.2	16.6	2163	4	US-07-454-543A-49
8	59.5	16.6	2161	1	US-08-455-543A-49
9	59.5	16.6	2161	1	US-08-455-543A-51
10	59.5	16.6	2161	2	US-08-223-305C-49
11	59.5	16.6	2161	2	US-08-223-305C-51
12	59.5	16.6	2161	2	US-08-311-363-2
13	58.5	16.3	374	2	US-08-628-692-51
14	56.5	15.2	3868	2	US-08-609-049A-22
15	54.5	15.2	3868	2	US-09-059-853-2
16	54.5	15.2	3868	2	US-09-059-853-2
17	54.5	15.0	242	2	US-08-626-352A-3
18	54.5	15.0	242	3	US-08-826-390-3
19	54.5	15.0	271	1	US-08-276-919-10
20	54.5	15.0	271	1	US-08-776-088-13
21	54.5	15.0	271	4	PCT-US95-09145A-13
22	54.5	15.0	271	4	US-08-609-088-15
23	54.5	15.0	325	1	US-08-609-088-15
24	54.5	15.0	325	1	US-08-276-919-4
25	54.5	15.0	325	1	US-08-276-919-13
26	54.5	15.0	325	1	US-08-276-919-13
27	54.5	15.0	325	1	US-08-776-088-4
28	54.5	15.0	325	4	US-08-776-088-18
					PCT-US95-09145A-4

29	54	15.0	325	4	PCT-US95-09145A-18
30	54	15.0	354	1	US-08-276-919-2
31	54	15.0	354	1	US-08-776-088-2
32	54	15.0	354	1	US-08-776-088-6
33	54	15.0	354	3	US-08-609-088-2
34	54	15.0	354	3	PCT-US95-09145A-2
35	54	15.0	354	4	PCT-US95-09145A-6
36	53.5	14.9	137	2	US-08-637-759B-34
37	53.5	14.9	137	3	US-08-637-759B-34
38	53	14.8	266	2	US-07-857-224B-40
39	53	14.8	266	2	US-07-857-224B-40
40	53	14.8	1289	2	US-08-690-473-2
41	52.5	14.6	312	3	US-08-414-625-2
42	52.5	14.6	312	3	US-08-414-625-2
43	52	14.5	338	2	US-08-674-149A-12
44	52	14.5	482	1	US-07-792-885A-1
45	52	14.5	482	1	US-08-142-439A-7

## ALIGNMENTS

RESULT 1  
US-08-609-103B-2  
: Sequence 2, Application US/0809103B  
: Patent No. 6133505  
: GENERAL INFORMATION:  
: TITLE: TRANSGENIC PLANTS AND SEEDS FOR OBTAINING SAME  
: INVENTOR: GRONBORN, BRUNO  
: ATTORNEY: PATENT FIRM, HATZEL  
: TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS FOR OBTAINING SAME  
: NUMBER OF SEQUENCES: 17  
: CORRESPONDENCE ADDRESS:  
: ADDRESSER: YOUNG & THOMPSON  
: STREET: 745 South 23rd Street  
: CITY: Arlington  
: STATE: Virginia  
: COUNTRY: U.S.A.  
: ZIP: 22202  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patent Release #1.0, Version #1.30  
: COUNTRY: U.S.A.  
: APPLICATION NUMBER: US/08/609,103B  
: FILING DATE: 17-MAR-1997  
: CLASSIFICATION: 800  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: FR 94.11040  
: FILING DATE: 15-SEP-1994  
: PRIOR APPLICATION DATA: NO PCT/FR95/01192  
: APPLICATION NUMBER: 15-SEP-1995  
: FILING DATE: 15-SEP-1995  
: ATTORNEY/AGENT INFORMATION:  
: NAME: PATCH, Andrew J.  
: REGISTRATION NUMBER: 32,925  
: REFERENCE/DOCKET NUMBER: USB94AL CNR TOM  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (703) 521-2297  
: TELEFAX: (703) 521-0573  
: TELEFAX: 248425 BMON  
: INFORMATION FOR SEQ ID NO: 2:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 359 amino acids  
: TYPE: amino acid  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
US-08-609-103B-2

Query Match 61.0% Score 219; DB 3; Length 359;  
Best Local Similarity 58.0%; Pred. No. 2.6e-21;  
Matches 40; Conservative 11; Mismatches 18; Indels 0; Gaps 0;

```
Oy 2 LVWGFEQVDRSGCGCOTSDNAAEALNASSKEEALQIIREKIPAAALQFPHNINSNLD 61
Db 111 LEMGFQIDGRSGGQOTANDAYAKAINAGSQALDVIRELAPDYVLFHFININSNLD 170
Oy 62 RFDKTPPE 70
Db 171 KVFQVPEAP 179

RESULT 2
US-08-809-103B-4
: Sequence 4, Application US/08809103B
: Patent No. 6133505
: GENERAL INFORMATION:
: APPLICANT: GRONENBORN, Bruno
: TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: YOUNG & THOMPSON
: STREET: 745 South 23rd Street
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA: US/08/809,103B
: FILING DATE: 17-MAR-1997
: PRIORITY INFORMATION:
: APPLICATION NUMBER: FR 94.11040
: FILING DATE: 15-SEP-1994
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA: US/08/809,103B
: FILING DATE: 17-MAR-1997
: PRIORITY INFORMATION:
: APPLICATION NUMBER: FR 94.11040
: FILING DATE: 15-SEP-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: PATCH, Andrew J.
: REGISTRATION NUMBER: 32,925
: REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 521-2297
: TELEFAX: (703) 685-0573
: TELEX: 248425 EMBON
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 359 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-809-103B-4

Query Match 61.0%; Score 219; DB 3; Length 359;
Best Local Similarity 56.0%; Pred. No. 2,6e-21;
Matches 40; Conservative 11; Mismatches 18; Indels 0;

Oy 2 LVWGFEQVDRSGCGCOTSDNAAEALNASSKEEALQIIREKIPAAALQFPHNINSNLD 61
Db 111 LEMGFQIDGRSGGQOTANDAYAKAINAGSQALDVIRELAPDYVLFHFININSNLD 170
Oy 62 RFDKTPPE 70
Db 171 KVFQVPEAP 179

RESULT 3
US-08-809-103B-6
: Sequence 4, Application US/08809103B
: Patent No. 6133505
: GENERAL INFORMATION:
: APPLICANT: GRONENBORN, Bruno
: TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: YOUNG & THOMPSON
: STREET: 745 South 23rd Street
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA: US/08/809,103B
: FILING DATE: 17-MAR-1997
: PRIORITY INFORMATION:
: APPLICATION NUMBER: FR 94.11040
: FILING DATE: 15-SEP-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: PATCH, Andrew J.
: REGISTRATION NUMBER: 32,925
: REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 521-2297
: TELEFAX: (703) 685-0573
: TELEX: 248425 EMBON
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 359 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-809-103B-6

Query Match 61.0%; Score 219; DB 3; Length 359;
Best Local Similarity 56.0%; Pred. No. 2,6e-21;
Matches 40; Conservative 11; Mismatches 18; Indels 0;

Oy 2 LVWGFEQVDRSGCGCOTSDNAAEALNASSKEEALQIIREKIPAAALQFPHNINSNLD 61
Db 111 LEMGFQIDGRSGGQOTANDAYAKAINAGSQALDVIRELAPDYVLFHFININSNLD 170
Oy 62 RFDKTPPE 70
Db 171 KVFQVPEAP 179

RESULT 4
US-08-809-103B-8
: Sequence 8, Application US/08809103B
: Patent No. 6133505
: GENERAL INFORMATION:
: APPLICANT: GRONENBORN, Bruno
: TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: YOUNG & THOMPSON
: STREET: 745 South 23rd Street
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA: US/08/809,103B
: FILING DATE: 17-MAR-1997
: PRIORITY INFORMATION:
: APPLICATION NUMBER: FR 94.11040
: FILING DATE: 15-SEP-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: PATCH, Andrew J.
: REGISTRATION NUMBER: 32,925
: REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 521-2297
: TELEFAX: (703) 685-0573
: TELEX: 248425 EMBON
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 359 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-809-103B-8

Query Match 61.0%; Score 219; DB 3; Length 359;
Best Local Similarity 56.0%; Pred. No. 2,6e-21;
Matches 40; Conservative 11; Mismatches 18; Indels 0;

Oy 2 LVWGFEQVDRSGCGCOTSDNAAEALNASSKEEALQIIREKIPAAALQFPHNINSNLD 61
Db 111 LEMGFQIDGRSGGQOTANDAYAKAINAGSQALDVIRELAPDYVLFHFININSNLD 170
Oy 62 RFDKTPPE 70
Db 171 KVFQVPEAP 179
```



```

1  COUNTRY: U.S.A.
2  ZIP: 22002
3  COMPUTER READABLE FORM: disk
4  FILETYPE: PC compatible
5  OPERATING SYSTEM: PC-DOS/MS-DOS
6  SOFTWARE: PatentIn Release #1.0, Version #1.30
7  CURRENT APPLICATION DATA:
8  APPLICATION NUMBER: US/08/809,103B
9  FILING DATE: 17-MAR-1997
10  INVENTOR: O'Connor, Johnson, and Kindness
11  PRIOR APPLICATION DATA:
12  APPLICATION NUMBER: PR 94.11040
13  FILING DATE: 15-SEP-1994
14  PRIOR APPLICATION DATA:
15  APPLICATION NUMBER: WO PCT/FR95/01192
16  FILING DATE: 15-SEP-1995
17  ATTORNEY/AGENT INFORMATION:
18  ATTORNEY/AGENT INFORMATION:
19  REGISTRATION NUMBER: 32,925
20  REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
21  TELECOMMUNICATION INFORMATION:
22  TELEPHONE: (703) 521-2297
23  TELEFAX: (703) 695-0573
24  TELEX: 248425 EMBON
25  INFORMATION FOR SEQ ID NO: 8:
26  SEQUENCE CHARACTERISTICS:
27  LENGTH: 1713 amino acids
28  TYPE: amino acid
29  TOPOLOGY: linear
30  MOLECULE TYPE: protein
31  US-08-809-103B-8
32
33  Query Match 61.0%; Score 219; DB 3; Length 359;
34  Best Local Similarity 51.0%; Pred No 12;
35  Matches 40; Conservative 11; Mismatches 18; Indels 0; Gaps 0;
36
37  Oy 2 LVKSEFQVDRSGRGCGQTSNDAAAEALN--ASSKEEALQ-IIREKIPAAALFQFINLNSNLD 61
38  |||||:|||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
39  Db 111 LEMGTQIDGRSGRGCGQTANDAYAKAINAGSKSQALQVKEIAPDYVLVLFHFNINSNLD 170
40
41  Oy 62 EIEKTPRP 70
42  |||||:|||||:
43  Db 171 KVFQVPPAP 179
44
45  RESULT 5
46  US-08-600-982-24
47  ? Sequence 24, Application US/08600982
48  ? GENERAL INFORMATION:
49  ? APPLICANT: Carter, William G.
50  ? APPLICANT: G11, Susanna C.
51  ? TITLE OF INVENTION: Epiligrin, an Epithelial Ligand for
52  ? TITLE OF INVENTION: Epiligrin, an Epithelial Ligand for
53  ? NUMBER OF SEQUENCES: 30
54  ? CORRESPONDENCE ADDRESSES:
55  ? ADDRESSEE: Christensen, O'Connor, Johnson, and Kindness
56  ? STREET: 1420 Fifth Avenue
57  ? CITY: Seattle
58  ? STATE: WA
59  ? COUNTRY: USA
60  ? ZIP: 98101-8100
61  ? COMPUTER READABLE FORM: disk
62  ? FILETYPE: PC compatible
63  ? OPERATING SYSTEM: PC-DOS/MS-DOS
64  ? SOFTWARE: PatentIn Release #1.0, Version #1.25
65  ? CURRENT APPLICATION DATA:
66  ? APPLICATION NUMBER: US/08/600,982
67  ? FILING DATE: 02-SEP-1994
68  ? CLASSIFICATION:
69  ? NAME: SUGI INFORMATION:
70  ? NAME: SUGI INFORMATION:
71  ? REGISTRATION NUMBER: 25,997
72  ? TELECOMMUNICATION INFORMATION:
73  ? TELEPHONE: (206) 682-8100
74  ? TELEFAX: (206) 224-0779
75  ? INFORMATION FOR SEQ ID NO: 24:
76  ? SEQUENCE CHARACTERISTICS:
77  ? LENGTH: 1713 amino acids
78  ? TYPE: amino acid
79  ? TOPOLOGY: linear
80  ? MOLECULE TYPE: protein
81  ? DESCRIPTION: E170 protein as translated from sequence of
82  ? DESCRIPTION: FIGURES 15A-15F, and as shown also in FIGURES 19A-19F
83  ? PCT-US94-10261A-24

```

```

1  CLASSIFICATION: 435
2  ATTORNEY/AGENT INFORMATION:
3  NAME: Shelton, Dennis K.
4  REGISTRATION NUMBER: 26,997
5  TELECOMMUNICATION INFORMATION:
6  TELEPHONE: (206) 682-8100
7  TELEFAX: (206) 224-0779
8  INFORMATION FOR SEQ ID NO: 24:
9  SEQUENCE CHARACTERISTICS:
10  LENGTH: 1713 amino acids
11  TYPE: amino acid
12  TOPOLOGY: linear
13  MOLECULE TYPE: protein
14  DESCRIPTION: E170 protein as translated from sequence
15  DESCRIPTION: of FIGURES 15A-15F, and as shown also in FIGURES
16  DESCRIPTION: 19A-19F
17  US-08-600-982-24
18
19  Query Match 17.7%; Score 63.5; DB 3; Length 1713;
20  Best Local Similarity 34.4%; Pred No 12;
21  Matches 21; Conservative 21; Mismatches 21; Indels 7; Gaps 3;
22
23  Oy 9 VDRSGRGCGQTSNDAAAEALN--ASSKEEALQ-IIREKIPAAALFQFINLNSNLD 65
24  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
25  Db 553 VDAATATAYENILNIAKAAEDAAANRAASASALQTVIKEDLPKRA---KTLSSNSDKLLN 608
26
27  Oy 66 K 66
28  |||||:|||||:
29  Db 609 E 609
30
31  RESULT 6
32  PCT-US94-10261A-24
33  ? Sequence 24, Application PC/TUS9410261A
34  ? GENERAL INFORMATION:
35  ? APPLICANT: Carter, William G.
36  ? APPLICANT: G11, Susanna A.
37  ? TITLE OF INVENTION: Epiligrin, an Epithelial Ligand for
38  ? TITLE OF INVENTION: Epiligrin, an Epithelial Ligand for
39  ? NUMBER OF SEQUENCES: 30
40  ? CORRESPONDENCE ADDRESSES:
41  ? ADDRESSEE: Christensen, O'Connor, Johnson, and Kindness
42  ? STREET: 1420 Fifth Avenue
43  ? CITY: Seattle
44  ? STATE: WA
45  ? COUNTRY: USA
46  ? ZIP: 98101-8100
47  ? COMPUTER READABLE FORM: disk
48  ? FILETYPE: PC compatible
49  ? OPERATING SYSTEM: PC-DOS/MS-DOS
50  ? SOFTWARE: PatentIn Release #1.0, Version #1.25
51  ? CURRENT APPLICATION DATA:
52  ? APPLICATION NUMBER: PCT/US94/10261A
53  ? FILING DATE: 02-SEP-1994
54  ? CLASSIFICATION:
55  ? NAME: SUGI INFORMATION:
56  ? NAME: SUGI INFORMATION:
57  ? REGISTRATION NUMBER: 25,997
58  ? TELECOMMUNICATION INFORMATION:
59  ? TELEPHONE: (206) 682-8100
60  ? TELEFAX: (206) 224-0779
61  ? INFORMATION FOR SEQ ID NO: 24:
62  ? SEQUENCE CHARACTERISTICS:
63  ? LENGTH: 1713 amino acids
64  ? TYPE: amino acid
65  ? TOPOLOGY: linear
66  ? MOLECULE TYPE: protein
67  ? DESCRIPTION: E170 protein as translated from sequence of
68  ? DESCRIPTION: FIGURES 15A-15F, and as shown also in FIGURES 19A-19F
69  ? PCT-US94-10261A-24

```



```

QY      4 WGEFVDGSRGARGCOSTS-----NDAAEALNASSK 34
DB      707 WNAVMYDGLMAYGTSSSGMVCIYFIPLFCGNVILLWFLAVDLADLAEASINTAQK 766

QY      35 EEALQIIREKI 45
DB      767 EEEKEKKRKI 777

RESULT 9
: Sequence 51, Application US/0845543A
: Patent No. 5792846
: GENERAL INFORMATION:
: APPLICANT: Michael, Michael
: APPLICANT: Ellis, Steven
: APPLICANT: Williams, Mark
: APPLICANT: Feldman, Daniel
: APPLICANT: McCue, Ann
: APPLICANT: Brenner, Robert
: TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
: METHODS FOR THEIR USE
: NUMBER OF SEQUENCES: 51
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Brown, Martin, Haller & McClain
: STREET: 1660 Union Street
: CITY: San Diego
: STATE: California
: COUNTRY: USA
: FILING DATE: April 4, 1994
: COMPUTER PROGRAM FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq Version 1.5
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/455.543A
: PRIORITY DATE: August 11, 1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/223.305
: FILING DATE: April 4, 1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/868.354
: FILING DATE: April 10, 1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/745.206
: FILING DATE: 16-AUG-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/620.250
: FILING DATE: 30-NOV-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/482.384
: FILING DATE: 20-FEB-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/603.751
: FILING DATE: 04-APR-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WO PCT/US89/01408
: FILING DATE: 04-APR-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/176.899
: FILING DATE: 04-MAR-1988
: ATTORNEY/AGENT INFORMATION:
: NAME: Seidman, Stephanie L.
: REGISTRATION NUMBER: 33,779
: REFERENCE/DOCKET NUMBER: 6362-S2517
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619)238-0099
: TELEFAX: (619)238-0052
: CONTACT PERSON: 51:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2161 amino acids
: TYPE: amino acid

```

```

: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: FRAGMENT TYPE: internal
US-08-455-543A-51

Query Match          16 %; Score 59.5; DB 1; Length 2161;
Best Local Similarity 28.2%; Pred No. 54;
Matches 20; Conservative 5; Mismatches 17; Indels 29; Gaps

QY      4  WGFDFVGGSGARGSCG-----NDAAEALNASSK 34
DB       707  WNAWVTGDTMAYGSSSMGVICVFIFILFCIGNTILLAVFLAVDNCADSLASINTAK 766
              ||| || | : |
QY      35  EEAQTITREKI 45
DB       767  EEAEKERKKI 777
              ||| : |||

US-RESULT 10
: Sequence 435 Application US/08223305C
: PCT NUMBER: 565244
: GENERAL INFORMATION:
: APPLICANT: Harbold, Michael
: APPLICANT: Ellis, Steven
: APPLICANT: Williams, Mark
: APPLICANT: Feldman, Daniel
: APPLICANT: McCue, Ann
: APPLICANT: Brenner, Robert
: TITLE OF INVENTION: CHANNELS AND ION CHANNEL COMPOSITIONS AND METHODS
: NUMBER OF SEQUENCES: 57
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Brown, Martin, Haller & McClain
: STREET: 1660 Union Street
: CITY: San Diego
: STATE: California
: COUNTRY: USA
: ZIP: 92101-2926
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq Version 1.5
: CURRENT APPLICATION DATA/US/08/223,305C
: PCT APPLICATION NUMBER: US 07/603,751
: FILING DATE: April 4, 1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/868,354
: FILING DATE: April 10, 1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/745,206
: FILING DATE: 15-AUG-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/620,250
: FILING DATE: 30-NOV-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/482,384
: FILING DATE: 20-FEB-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/603,751
: FILING DATE: 04-APR-1988
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WO PCT/US89/01408
: FILING DATE: 04-APR-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/176,899
: FILING DATE: 04-APR-1988
: ATTORNEY/AGENT INFORMATION:
: NAME: Schindler, Stephen L.
: REGISTRATION NUMBER: 33,779

```

REFERENCE/DOCKET NUMBER: 52516 (P519739)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619)238-0999  
FILING DATE: 04-APR-1999  
INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2161 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORGANISM: Homo sapiens  
US-08-223-305C-49

Query Match 16.6%; Score 59.5; DB 2; Length 2161;  
Best Local Similarity 28.2%; Pred. No. 54;  
Matches 20; Conservative 5; Mismatches 17; Indels 29; Gaps 1;  
Oy 4 WGEFQVGRSARGCQTS-----NDAAAEALNASSK 34  
Db 707 WNAVYDGTMAVGPSGGMVICVFTILFCGNYILLNVFLAVDNLADESINATQK 766

Oy 35 BEALQIREKI 45  
Db 767 BEAEKERKKI 777

RESULT 11  
US-08-223-305C-51  
Sequence 51, Application US/08223305C  
Patent No. 5851824  
GENERAL INFORMATION:  
APPLICANT: Harpold, Michael  
APPLICANT: Ellis, Steven  
APPLICANT: Williams, Mark  
APPLICANT: Feldman, Daniel  
APPLICANT: McCue, Ann  
APPLICANT: Brenner, Robert  
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND  
NUMBER OF SEQUENCES: 57  
COMBINED ADDRESS:  
STREET: 1660 Union Street  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-2926  
COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
FILING DATE: April 4, 1994  
APPLICATION NUMBER: 07/868,354  
FILING DATE: April 10, 1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/745,206  
FILING DATE: 15-AUG-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/620,250  
FILING DATE: 15-AUG-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/450,350  
FILING DATE: 20-FEB-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/603,751  
FILING DATE: 04-APR-1989  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US89/01408  
FILING DATE: 04-APR-1989  
PRIOR APPLICATION DATA:  
PCT NUMBER: US 07/176,899  
PCT DATE: 04-APR-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Seidman, Stephanie L.  
REGISTRATION NUMBER: 33,779  
REFERENCE/DOCKET NUMBER: 52516 (P519739)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619)238-0999  
FILING DATE: 04-APR-1999  
INFORMATION FOR SEQ ID NO: 51:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2161 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORGANISM: Homo sapiens  
US-08-223-305C-51

Query Match 16.6%; Score 59.5; DB 2; Length 2161;  
Best Local Similarity 28.2%; Pred. No. 54;  
Matches 20; Conservative 5; Mismatches 17; Indels 29; Gaps 1;  
Oy 4 WGEFQVGRSARGCQTS-----NDAAAEALNASSK 34  
Db 707 WNAVYDGTMAVGPSGGMVICVFTILFCGNYILLNVFLAVDNLADESINATQK 766

Oy 35 BEALQIREKI 45  
Db 767 BEAEKERKKI 777

RESULT 12  
US-08-311-363-2  
Sequence 2, Application US/08311363  
Patent No. 5876958  
GENERAL INFORMATION:  
APPLICANT: Harpold, Michael  
APPLICANT: Ellis, Steven  
APPLICANT: Williams, Mark  
APPLICANT: Feldman, Daniel  
APPLICANT: McCue, Ann  
APPLICANT: Brenner, Robert  
TITLE OF INVENTION: Human Calcium Channel Compositions and  
NUMBER OF SEQUENCES: 32  
COMBINED ADDRESS:  
STREET: 1660 Union Street  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-2926  
COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/311,363  
FILING DATE: 15-AUG-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/745,206  
FILING DATE: 15-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Seidman, Stephanie L.  
REGISTRATION NUMBER: 33,779  
REFERENCE/DOCKET NUMBER: 6362-51506  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (619)238-0999  
 TELEFAX: (619)238-0062  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 216 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-311-363-2

Query Match 16.5%; Score 59.5; DB 2; Length 2161;  
 Best Local Similarity 2.24%; Pred. No. 54;  
 Matches 20; Conservative 5; Mismatches 17; Indels 29; Gaps 1;  
 Qy 4 WCEFDVGRSARGCQTS-----NDAAALNASSK 34  
 Db 707 WNAVYDGIWAVGGSSGMCIVFIILFIGNYILNVLAIADVNLDAESLNTAOK 766  
 Qy 35 EALQIREKI 45  
 Db 767 EAEKERKKI 777

RESULT 13  
 US-08-928-692-51  
 Sequence 51, Application US/08928692  
 Patent No. 5958721  
 GENERAL INFORMATION:  
 APPLICANT: Brady, Howard  
 APPLICANT: Yaver, Deborah S.  
 APPLICANT: Lamsa, Michael  
 APPLICANT: Hansen, Kim  
 TITLE OF INVENTION: Methods for Modifying the Production of  
 TITLE OF INVENTION: a Polypeptide  
 NUMBER OF SEQUENCES: 80  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: No. 59587270 No. 5958727 disk of No. 5958727 America, Inc.  
 STREET: 405 Lexington Avenue  
 CITY: New York  
 STATE: NY  
 COUNTRY: USA  
 ZIP: 10174  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5 inch floppy diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/928,692  
 FILING DATE: 12-SEPT-1997  
 PUBLICATION DATE: 05-SEP-1999  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Lambiris, Elias J.  
 REGISTRATION NUMBER: 33,728  
 REFERENCE/DOCKET NUMBER: 4944,200-US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-878-9655  
 TELEFAX: 212-878-0123  
 INFORMATION FOR SEQ ID NO: 51:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 374 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-928-692-51

Query Match 16.3%; Score 58.5; DB 2; Length 374;  
 Best Local Similarity 45.5%; Pred. No. 7.2;  
 Matches 15; Conservative 5; Mismatches 10; Indels 3; Gaps 1;

Qy 22 NDAAALNASSKE---EALQIREKIPAAALF 51  
 Db 181 SDAQREALTASAKAAGLEVQLIHEPVAALAY 213  
 RESULT 14  
 US-08-609-049A-22  
 Sequence 22, Application US/08609049A  
 Patent No. 5948664  
 GENERAL INFORMATION:  
 APPLICANT: Williams, Lewis T.  
 APPLICANT: Molz, Lisa  
 APPLICANT: Molz, Lisa  
 TITLE OF INVENTION: 3-Kinase Polypeptides  
 NUMBER OF SEQUENCES: 32  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Townsend and Townsend and Crew LLP  
 STREET: Two Embarcadero Center, 8th floor  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94111-3834  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/609,049A  
 FILING DATE: 29-FEB-1996  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Dow, Karen B.  
 REGISTRATION NUMBER: 29,684  
 REFERENCE/DOCKET NUMBER: 2307K-0637000S  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-326-2400  
 TELEFAX: 415-326-2422  
 INFORMATION FOR SEQ ID NO: 22:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 171 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MEDIUM TYPE: Peptide  
 US-08-609-049A-22

Query Match 15.6%; Score 56; DB 2; Length 171;  
 Best Local Similarity 23.4%; Pred. No. 5.5;  
 Matches 18; Conservative 14; Mismatches 33; Indels 12; Gaps 2;  
 Qy 2 LVNKEFDVGRSARGCQTSNDAA--PALNASSKEALQIREKTPAAAL-----50  
 Db 24 LNKKEAGLDIRMLPGCLATGDSGLIEVYVSTETIADIQLNSSNVAATAAFNKKALLN 83  
 Qy 51 -FOPHNLNSLDRIFFDK 66  
 Db 84 WKEKNSGGDLDRATEE 100

RESULT 15  
 US-08-876-991-2  
 Sequence 2, Application US/08876991  
 Patent No. 5925360  
 GENERAL INFORMATION:  
 APPLICANT: Gregor Meyers, Tillmann R menapf,  
 APPLICANT: Heitz-J iden Thiel  
 APPLICANT: Weyer, Hans-Joachim  
 NUMBER OF SEQUENCES: 13  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Organon Teknika Corporation  
 ADDRESSEE: Biotechnology Research Institute

```

STREET: 1330-A Piccard Drive
CITY: Rockville
STATE: Maryland
COUNTRY: U.S.A.
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/876,991
  FILING DATE: 16-JUN-1997
  CLASSIFICATION: 424
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US/08/747,577
  FILING DATE:
  APPLICATION NUMBER: US/08/650,584
  FILING DATE:
  APPLICATION NUMBER: US/08/469,702
  FILING DATE:
  APPLICATION NUMBER: US/08/123,596
  FILING DATE:
  APPLICATION NUMBER: 07/797,554
  FILING DATE: 22-NOV-1991
  APPLICATION NUMBER: US 07/494,991
  FILING DATE: 16-MAR-1990
  CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
  NAME: William M. Blackstone
  REGISTRATION NUMBER: 29,772
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (301) 258-5200
  FACSIMILE: (301) 258-5200
  INQUIRY OR CORRESPONDENCE TO: 2
SEQUENCE CHARACTERISTICS:
  LENGTH: 3899 amino acids
  TYPE: amino acid
  TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-876-991-2

```

```

Query Match      15.74; Score 54.5; D8 2; Length 3898;
Best Local Similarity 21.4%; Pred No. 5.5c+02;
Matches, 18; Conservative 15; Mismatches 34; Indels 17; Gaps 4;

QY 1 TLVWGEF---QVDRSGARGCQTSNDAAEALNASSKEEALQIINREKIPAAALFQ--FH 54
Db 716 ITTWKFSHGQLQDDVTVKAVC-TAGSFVTVTAUNVSRVYLAASHRRALPTVTFELFD 774
QY 55 NLNSLDRP-----FDKTP 68
Db 775 GTNPAIEEMDDDFGFLGCPDTPSP 798

```

Search completed: February 3, 2001, 02:19:45  
 Job time: 1707 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:24:26 ; Search time 83.07 seconds  
(without alignments)  
27.213 Million cell updates/sec

Title: US-09-289-346a-5

Perfect score: 359

Sequence: 1 TLVWGEFVDGRSGGQCT.....QPFHNSNLDRIKDTPPEP 70

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	343	95.5	322	1 VAL1_TGMV	P03567 tomato gold
2	247	69.8	361	1 VAL1_PYYV	P27258 potato yell
3	243	67.7	358	1 VAL1_CLVN	P14972 cassava lat
4	232	64.6	362	1 VAL1_TYLCV	P36279 tomato yell
5	226	63.0	359	1 VAL1_TYLCV	P38609 tomato yell
6	224	62.4	349	1 VAL1_PHYV	P06923 pepper huas
7	219	61.0	359	1 VAL1_TYLCV	P27260 tomato yell
8	216	60.2	353	1 VAL1_BGMV	P05175 bean golden
9	216	59.2	361	1 VAL1_TMOX	P06557 tomato mott
10	212	59.1	361	1 VAL1_PHYV	P14977 bellonion mo
11	212	59.1	358	1 VAL1_PHYV	P14977 bellonion mo
12	212	59.1	358	1 VAL1_PHYV	P14977 bellonion mo
13	203	56.5	357	1 VAL1_TYLCV	P27259 tomato yell
14	118	32.9	347	1 VAL1_SICV	P29048 squash leaf
15	65	18.1	630	1 VAL1_BPV66	P08057 human papil
16	63.5	17.7	1713	1 LMA3_CANAL	P16787 homo sapien
17	62.5	17.4	428	1 GRAL1_CANAL	P28858 candida alb
18	60.5	16.9	1070	1 P311_HUMAN	Q05112 human papil
19	60.5	16.9	1070	1 P311_HUMAN	Q05112 human papil
20	59.5	16.6	705	1 Y006_RICPR	Q92406 richestia
21	59.5	16.6	1053	1 UBA3_MESAU	P31252 triticum ae
22	59.5	16.6	1610	1 CCAD_HUMAN	Q99244 mesocricetu
23	59.5	16.6	2161	1 CCAD_HUMAN	Q01668 homo sapien
24	59.5	16.6	2203	1 CCAD_RAT	P27732 rattus norv
25	59	16.4	703	1 UGS2_RAT	P71625 rattus norv
26	58	16.2	1852	1 CAS_CITPCA	P22316 cyprinus ca
27	58	16.2	1852	1 CAS_CITPCA	P22316 cyprinus ca
28	58	16.2	1608	1 HLXA_SERNR	P55320 zebrafilmu s
29	57.5	16.0	367	1 LHX4_MGUSE	P53776 mus musculu
30	57	15.9	355	1 CRTB_RIOSE	P54905 rhodobacter
31	57	15.9	524	1 MPPA_RAT	P20069 rattus norv
32	57	15.9	874	1 SLAP_BACLI	P49052 bacillus li
33	56	15.6	129	1 RK12_PORPU	P51339 porphyra pu

34	56	15.6	329	1 TC3A_CABEL	P34257 caenorhabdi
35	55.5	15.5	266	1 ETFB_MYCLE	O33095 mycobacteri
36	55.5	15.5	266	1 ETFB_MYCTU	O33276 mycobacteri
37	55	15.5	2190	1 CCAD_CHICK	O73700 gallus gall
38	55	15.3	268	1 TL18_RABIT	P14922 actinoplasm
39	55	15.3	268	1 TL18_RABIT	P14922 actinoplasm
40	55	15.3	362	1 PGL2_ASPTU	P19805 aspergillus
41	55	15.3	408	1 CINA_THEME	O965x1 thermotoga
42	55	15.3	433	1 UFGO_SOLME	Q43641 solanum mel
43	55	15.3	480	1 STE_HAEN	P43818 haemophilus
44	55	15.3	525	1 MPPA_HUMAN	Q10713 homo sapien
45	55	15.3	576	1 HEMA_SENDR	P03425 sendai viru

## ALIGNMENTS

RESULT	1
VAL1_TGMV	TGMV
VAL1_PYYV	PYYV
ALC	STANDARD; PRT: 352 AA.
P03567	
DT	21-JUL-1986 (Rel. 01, Created)
DT	21-JUL-1986 (Rel. 01, Last sequence update)
DT	01-JUN-1994 (Rel. 29, Last annotation update)
DE	AL1 PROTEIN.
GN	ACL.
OS	Tomato golden mosaic virus (TGMV).
OC	Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RA	Sequence from N.A.
RP	Sequence from N.A.
RT	Hamilton W.D.O., Stein V.E., Gotts R.H.A., Buck K.W.;
RT	*Complete nucleotide sequence of the infectious cloned DNA components
RT	of tomato golden mosaic virus: potential coding regions and regulatory
RT	sequences.;
RL	EMBL J. 3:2197-2205(1984).
CC	--- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed, usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	-- send an email to: license@isb-sib.ch).
CC	EMBL: K02029; -: NOT_ANNOTATED_CDS.
DR	PIR: A04170; Q0CVLL.
DR	INTERPRO: IPR001191; -:
DR	INTERPRO: IPR001301; -:
DR	PFAM: PF00799; Gemini1_A1; 1.
DR	PRINTS: PR00227; GEMCONTALL1.
DR	SMART: SM00226; GEMCONTALL1.
KW	ATP-binding; 223 230 ATP (POTENTIAL).
FT	NP_BIND 223 230
SQ	SEQUENCE 352 AA: C33C938E9644B4A4 CRC64:

Query Match	95.5%	Score 343;	DB 1;	Length 352;
Best Local Similarity	95.7%	Pred. No. 9.5e-32;		
Matches	67;	Conservative	0;	Mismatches 3;
				Indels 0;
				Gaps 0;
Oy	1	TLVWGEFVDGRSGGQCTSDAAALAAALASKEALQIIRKIPAAALQFPHNSNL	60	
Db	111	TLVWGEFVDGRSGGQCTSDAAALAAALASKEALQIIRKIPKYLQFPHNSNL	170	
Oy	61	DRIFDKTPPEP 70		
Db	171	DRIFDKTPPEP 180		
RESULT	2			
VAL1_PYYV				
ID	VAL1_PYYV	STANDARD; PRT: 361 AA.		

```

AC DT 01-AUG-1992 (Rel. 23, Created)
AD DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE AL1 PROTEIN
OC Potalo yellow mosaic virus (Isolate Venezuela).
OC Viruses: ssDNA viruses: Geminiviridae; Begomovirus.
RN [1]
RX MEDLINE:91311403; PubMed:1856690;
RA Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.:
RL "The nucleotide sequence of the infectious cloned DNA components of
RT potyvirus, 'Virus 72-1515-1520(1991).
RL Gen. Virol. 72:1515-1520(1991).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC EMBL: D00940; BAA00782.1; -
DR PIR: J03364; Q0CVPT.
DR INTERPRO: IPR001191; -
DR PFAM: PF00799; GeminL.A1; 1.
DR PRINTS: PR00227; GEMCOATALL.
DR PRINTS: PR00228; GEMCOATCLV1.
KW ATP-binding.
KW NP-BIND.
FT SEQUENCE 361 AA: 40650 MW: 5627A33BF1264383 CRC64;
SQ SEQUENCE 361 AA: 40650 MW: 5627A33BF1264383 CRC64;

Query Match 68.84; Score 247; DB 1; Length 361;
Best Local Similarity 68.1%; Pred. No. 7, 8e-21;
Matches 47; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

Qy 1 TLVWGEFQVDSARGGQTSNDAAAFALNASKEEALQIREKIPAAALFQFHNLNSL 60
Db 110 TVMGQFQIDGSRAGGQTVNDAAAFALNASKEEALQIREKIPEREYFQFHNLNSL 169
Qy 61 DRIFDKPTE 69
Db 170 DRIFDKAPE 178

RESULT 3
ID VAIL CLVK STANDARD; PRT: 358 AA.
AC P14982;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DE 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN (40.4 KDA PROTEIN).
OS Cassava latent virus (strain West Konyan 844).
OC Viruses: ssDNA viruses, Geminiviridae, Begomovirus.
RN [1]
RX Stanley J., Gay M.R.:
RA "Nucleotide sequence of cassava latent virus DNA.";
RL Nature 301:260-262(1983)
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
CC EMBL: J02057; -; NOT_ANNOTATED_CDS.
DR INTERPRO: IPR001191; -
DR PFAM: PF00799; GeminL.A1; 1.
DR PRINTS: PR00227; GEMCOATALL.
DR PRINTS: PR00228; GEMCOATCLV1.
KW ATP-binding.
KW NP-BIND.
FT SEQUENCE 358 AA: 40345 MW: ED173E753E95D69 CRC64;
SQ SEQUENCE 358 AA: 40345 MW: ED173E753E95D69 CRC64;

Query Match 67.74; Score 243; DB 1; Length 358;
Best Local Similarity 64.3%; Pred. No. 2, 2e-20;
Matches 45; Conservative 11; Mismatches 14; Indels 0; Gaps 0;

Qy 1 TLVWGEFQVDSARGGQTSNDAAAFALNASKEEALQIREKIPAAALFQFHNLNSL 60
Db 109 TVMGQFQIDGSRAGGQTSNDAAAFALNASKEEALQIREKIPEREYFQFHNLNSL 168
Qy 61 DRIFDKPTE 70
Db 169 DRIFDKPEAP 178

RESULT 4
ID VAIL CLVK STANDARD; PRT: 358 AA.
AC P14972;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DE 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN (40.4 KDA PROTEIN).
OS Cassava latent virus (strain Nigerian).
OC Viruses: ssDNA viruses, Geminiviridae, Begomovirus.
RN [1]
RX MEDLINE:90174930; PubMed:2308831;
RA Morris B., Coates L., Lowe S., Richardson K., Eddy P.:
RT "Nucleotide sequence of the infectious cloned DNA components of
CC African cassava mosaic virus (Nigerian strain)."
CC Nucleotide sequence of the infectious cloned DNA components of
CC African cassava mosaic virus (Nigerian strain).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC EMBL: X17095; CAA34953.1; -
DR PIR: S07594; S07594.
DR INTERPRO: IPR001191; -
DR PFAM: PF001301; -
DR PRINTS: PR00227; GEMCOATALL.
DR PRINTS: PR00228; GEMCOATCLV1.
KW ATP-binding.
KW NP-BIND.
FT SEQUENCE 358 AA: 40435 MW: 1DB16B80CB2D5E2C CRC64;
SQ SEQUENCE 358 AA: 40435 MW: 1DB16B80CB2D5E2C CRC64;

Query Match 67.74; Score 243; DB 1; Length 358;
Best Local Similarity 64.3%; Pred. No. 2, 2e-20;
Matches 45; Conservative 11; Mismatches 14; Indels 0; Gaps 0;

Qy 1 TLVWGEFQVDSARGGQTSNDAAAFALNASKEEALQIREKIPAAALFQFHNLNSL 60
Db 109 TVMGQFQIDGSRAGGQTSNDAAAFALNASKEEALQIREKIPEREYFQFHNLNSL 168

```



```

Oy 61 DRIFDXTPEP 70
Db 169 DRIFQFPAP 178

RESULT 5
ID VAL1.TYLCA STANDARD; PRT; 362 AA.
AC P36279;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN (C1 PROTEIN).
GN C1.
OS Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP SEQUENCE FROM N.A.
RA DRY I.B. Gliden, J.P. Erskine L.R., Mullineaux P.M., Rezanian M.A.;
RT "Nucleotide sequence and genome organization of tomato leaf curl
geminivirus.";
RL J. Gen. Virol. 74:147-151(1993).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
DR EMBL; Z25751; CAAB1026.1;
DR PIR; S39211; S39211.
DR INTERPRO; IPR001191;
DR PIR; P36279; GEMINIVIR;
DR PFAM; PF00759; GEMIN1;
DR PRINTS; PR00228; GEMCONTAL1.
KW ATP-binding. 221 228 ATP (POTENTIAL).
FT NP_BIND 221 228 ATP (POTENTIAL).
SQ SEQUENCE 362 AA; 41197 MW; 3435718464/04098 CRC64;

Query Match 64.6%; Score 232; DB 1; Length 362;
Best Local Similarity 55.3%; Pred. No. 4e-19;
Matches 47; Conservative 8; Mismatches 14; Indels 16; Gaps 1;

Oy 1 TLVGEFQVDRSGAGCGGQTSNDAAEALNASSKEALQIIREKIPAAALFQHNLSNI 60
Db 110 TLVGEFQIDGRSGAGCGGQTSNDAAEALNASSKEALQIIREKIPAAALFQHNLSNI 169

Oy 61 DRI
Db 170 DRIFTPLVYSPFLSSSFORVPE 194

RESULT 6
ID VAL1.TYLCA STANDARD; PRT; 359 AA.
AC P36609;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE AL1 PROTEIN (C1 PROTEIN).
GN C1.
OS Tomato yellow leaf curl virus (strain Murcia) (TYLCV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-94256836; PubMed-8198442;
RT "Complete nucleotide sequence of pepper huasteco virus: analysis and
comparison with bipartite geminiviruses.";
RL J. Gen. Virol. 74:2225-2231(1993).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
DR EMBL; X70418; CAA49856.1;
DR PIR; S31875; S31875.
DR PIR; JQ2300; JQ2300.
DR INTERPRO; IPR001191;

```

```

RA Noris E., Hidalgo E., Accotto G., Moriones E.;
RT "High similarity among the tomato yellow leaf curl virus isolates
from the west Mediterranean basin: the nucleotide sequence of an
infectious clone from Spain";
RL Arch. Virol. 135:103-104.
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
DR EMBL; Z25751; CAAB1026.1;
DR PIR; S39211; S39211.
DR INTERPRO; IPR001191;
DR PIR; P39211; GEMINIVIR;
DR PFAM; PF00759; GEMIN1;
DR PRINTS; PR00228; GEMCONTAL1.
KW ATP-binding. 221 228 ATP (POTENTIAL).
FT NP_BIND 221 228 ATP (POTENTIAL).
SQ SEQUENCE 359 AA; 41065 MW; 2D170A51EF80A3EC CRC64;

Query Match 63.0%; Score 226; DB 1; Length 359;
Best Local Similarity 60.9%; Pred. No. 1.9e-18;
Matches 42; Conservative 9; Mismatches 18; Indels 0; Gaps 0;

Oy 2 LVKGEFQVDRSGAGCGGQTSNDAAEALNASSKEALQIIREKIPAAALFQHNLSNI 61
Db 111 LVKGEFQIDGRSGAGCGGQTSNDAAEALNASSKEALQIIREKIPAAALFQHNLSNI 170

Oy 62 RIFQKTPPEP 70
Db 171 RVFQVFPAP 179

RESULT 7
ID VAL1.PHUV STANDARD; PRT; 349 AA.
AC P06378;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN.
GN AL1.
OS Pepper huasteco virus (PHV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP SEQUENCE FROM N.A.
RA Torres-Pacheco I., Garzon-Riznado J.A., Herrera-Estrella L.,
RT "Complete nucleotide sequence of pepper huasteco virus: analysis and
comparison with bipartite geminiviruses.";
RL J. Gen. Virol. 74:2225-2231(1993).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
DR EMBL; X70418; CAA49856.1;
DR PIR; S31875; S31875.
DR PIR; JQ2300; JQ2300.
DR INTERPRO; IPR001191;

```

```

DR INTERPRO: IPR001301;
DR PFAM: PF00799; Gemini_ALI.1.
DR PRINTS: PR00227; GEMCOATL.VL.
DR KEGG: K01210; GEMCOATL.VL.
KW ATP-BINDING; 221 228 ATP (BY SIMILARITY).
FT NP-BIND 349 AA; 39722 MW; D5F4E76CD56370F4 CRC64;
SQ SEQUENCE 349 AA; 39722 MW; D5F4E76CD56370F4 CRC64;

Query Match 62.4%; Score 224; DB 1; Length 349;
Best Local Similarity 61.4%; Pred. No. 3.1e-18;
Matches 43; Conservative 11; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLMWGEFQVDSRGSGCQTNDAAALNASSKEALQIIRKIPAAALFQFHNLNSL 60
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 110 TLMWGEFQVDSRGSGCQTNDAAALNASSKEALQIIRKIPAAALFQFHNLNSL 169
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 DRIFDKTPPEP 70
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 170 NRIFQTPPEP 179
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 8
VAL1_TYLCM
ID VAL1_TYLCM STANDARD; PRT; 359 AA.
AC P27260;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-OCT-1996 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN (C1 PROTEIN).
GN CL.
OS Tomato yellow leaf curl virus (strain Marmande) (TYLCV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN SEQUENCE FROM N.A.
RP Knapton B.; Bendall M.; Matzelt V.; Accotto G.P.; Crespi S.;
RA Knapton B.; Bendall M.; Matzelt V.; Accotto G.P.; Crespi S.;
RA Gronenborn B.;
RT "Tomato yellow leaf curl virus from Sardinia is a
   whitefly-transmitted monopartite geminivirus."
RL Nucleic Acids Res. 19:6763-6769(1991).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
DR EMBL: X61153; CAA43466.1;
DR PIR: S22593; S22593;
DR INTERPRO: IPR001191;
DR INTERPRO: IPR001301;
DR PFAM: PF00799; Gemini_ALI.1.
DR PRINTS: PR00227; GEMCOATL.VL.
DR KEGG: K01210; GEMCOATL.VL.
KW ATP-BINDING; 220 227 ATP (POTENTIAL).
FT NP-BIND 359 AA; 40733 MW; 9717940C93FEA7 CRC64;
SQ SEQUENCE 359 AA; 40733 MW; 9717940C93FEA7 CRC64;

```

```

Query Match 61.0%; Score 219; DB 1; Length 359;
Best Local Similarity 58.0%; Pred. No. 1.2e-17;
Matches 40; Conservative 11; Mismatches 18; Indels 0; Gaps 0;

QY 2 TLMWGEFQVDSRGSGCQTNDAAALNASSKEALQIIRKIPAAALFQFHNLNSL 61
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 111 TLMWGEFQVDSRGSGCQTNDAAALNASSKEALQIIRKIPAAALFQFHNLNSL 170
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 62 DRIFDKTPPEP 70
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db 171 KVFQVFPAP 179
RESULT 9
VAL1_BGVV
ID VAL1_BGVV STANDARD; PRT; 353 AA.
AC P05175;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE AL1 PROTEIN (40.2 KDA PROTEIN).
GN AC1. golden mosaic virus
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN SEQUENCE FROM N.A.
RP Howarth A.J.; Caton J.; Bossert M.; Goodman R.M.;
RT "Nucleotide sequence of bean golden mosaic R.M.;"
RT Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
DR EMBL: M10070; AAA46318.1;
DR INTERPRO: IPR001191;
DR INTERPRO: IPR001301;
DR PFAM: PF00799; Gemini_ALI.1.
DR PRINTS: PR00227; GEMCOATL.VL.
DR KEGG: K01210; GEMCOATL.VL.
KW ATP-BINDING; 222 229 ATP (POTENTIAL).
FT NP-BIND 353 AA; 40190 MW; 80FA79DF6029A34 CRC64;
SQ SEQUENCE 353 AA; 40190 MW; 80FA79DF6029A34 CRC64;

Query Match 60.2%; Score 216; DB 1; Length 353;
Best Local Similarity 60.0%; Pred. No. 3e-17;
Matches 42; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLMWGEFQVDSRGSGCQTNDAAALNASSKEALQIIRKIPAAALFQFHNLNSL 60
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 110 TLMWGEFQVDSRGSGCQTNDAAALNASSKEALQIIRKIPAAALFQFHNLNSL 169
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 DRIFDKTPPEP 70
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 170 ERIFKVPED 179
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 10
VAL1_TMOV
ID VAL1_TMOV STANDARD; PRT; 361 AA.
AC Q06657; 1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN.
GN ALL.
OS Tomato mottle virus (isolate Florida) (TMOV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN SEQUENCE FROM N.A.
RP MEDLINE: 93107858; Pubmed: 1469361;
RA Abouzaid A.M.; Polston J.E.; Hiebert E.;
RT "The nucleotide sequence of tomato mottle virus, a new geminivirus
   isolated from tomatoes in Florida."
RL J. Gen. Virol. 73:3225-3229(1992).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as the content is not modified
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L14460; AAC32414.1; -
DR EMBL: Q07670; -
DR INTERPRO: IP0001301; -
DR PFAM: PF00759; Gemini_AL1.1;
DR PRINTS: PR00227; GEMCATAL1.
DR PRINTS: PR00228; GEMCATCLVL1.
KW ATP-binding. 222 229 ATP (BY SIMILARITY).
FT NP_BIND 222 229
SQ SEQUENCE 361 AA: 40516 MW; 8130B65CEAC6950 CRC64;
-----
Query Match 60.24; Score 216; DB 1; Length 361;
Best Local Similarity 57.18; Pred. No. 2.6e-17;
Matches 40; Conservative 14; Mismatches 16; Indels 0; Gaps 0;
OY 1 TLVWGEFVQDSRGSGCOTSDAAAEALNASKEEAQLIIREKTPAAALQFHNLSNL 60
DB 110 TLMWGEFVQDSRGSGCOTSDAAAEALNASKEEAQLIIREKTPAAALQFHNLSNL 169
OY 61 DRIFDTPPEP 70
DB 170 ERIFAKAPEP 179
-----
RESULT 11
VALL_ARMYW STANDARD: PRT: 355 AA.
AC P21947:
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
CN ALL PROTEIN.
OS Abutilon mosaic virus (isolate West India).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP MEDLINE-91020984; PubMed-2219703;
RA Frischmuth T., Zimmat G., Jeske H.;
RT The nucleotide sequence of abutilon mosaic virus reveals prokaryotic
BL Vir190 178-461-668(1990).
BL -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X15983; -; NOT_ANNOTATED_CDS.
DR PIR: A36214; QOCVW1.
DR INTERPRO: IP0001191; -
DR PFAM: PF00759; Gemini_AL1.1;
DR PRINTS: PR00227; GEMCATAL1.
DR PRINTS: PR00228; GEMCATCLVL1.
KW ATP-binding. 221 228 ATP (POTENTIAL).
FT NP_BIND 221 228
SQ SEQUENCE 355 AA: 40257 MW; 16A2CABA63251B95 CRC64;
-----

```

```

Query Match 59.94; Score 215; DB 1; Length 355;
Best Local Similarity 58.64; Pred. No. 3.3e-17;
Matches 41; Conservative 11; Mismatches 18; Indels 0; Gaps 0;
OY 1 TLVWGEFVQDSRGSGCOTSDAAAEALNASKEEAQLIIREKTPAAALQFHNLSNL 60
DB 110 TLMWGEFVQDSRGSGCOTSDAAAEALNASKEEAQLIIREKTPAAALQFHNLSNL 169
OY 61 DRIFDTPPEP 70
DB 170 ERIFAKAPEP 179
-----
RESULT 12
VALL_BCTV STANDARD: PRT: 358 AA.
AC P14991:
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE ALI PROTEIN (40.8 KDA PROTEIN).
OS Best curly top virus (BCTV).
OC Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
RN [1]
RP Stanley J., Markham P.G., Callis R.J., Plimmer M.S.;
RA The nucleotide sequence of an infectious clone of the geminivirus
RT best curly top virus.
BL EMBL J. 5:1761-1767(1986).
BL -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X04144; -; NOT_ANNOTATED_CDS.
DR INTERPRO: IP0001191; -
DR PFAM: PF00759; Gemini_AL1.1;
DR PRINTS: PR00227; GEMCATAL1.
DR PRINTS: PR00228; GEMCATCLVL1.
KW ATP-binding. 222 229 ATP (POTENTIAL).
FT NP_BIND 222 229
SQ SEQUENCE 358 AA: 40889 MW; 39A45FE3C0B9C333 CRC64;
-----
Query Match 59.18; Score 212; DB 1; Length 358;
Best Local Similarity 55.74; Pred. No. 7.3e-17;
Matches 39; Conservative 15; Mismatches 4; Indels 0; Gaps 0;
OY 1 TLVWGEFVQDSRGSGCOTSDAAAEALNASKEEAQLIIREKTPAAALQFHNLSNL 60
DB 110 TLMWGEFVQDSRGSGCOTSDAAAEALNASKEEAQLIIREKTPAAALQFHNLSNL 169
OY 61 DRIFDTPPEP 70
DB 170 OKIFOREPDP 179
-----
RESULT 13
VALL_TYICV STANDARD: PRT: 357 AA.
AC P27259:
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE ALI PROTEIN (C1 PROTEIN).
GN C1.

```

```

OS Tomato yellow leaf curl virus (TYLCV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
EX SEQUENCE FROM N.A.
RA MEDLINE:91024070; PubMed:1926771;
RA Navot N., Pichersky E., Zeidan M., Zamil D., Czosnek H.;
RT "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
RT with a single genomic component.";
RL Virology 185:151-161(1991).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: X15556; GenBank: U00001;
CC PIR: Q15556; CCNC1368.1; -.
CC PIR: Q15556; CCNC1368.1; -.
DR INTERPRO: IPR001191; -.
DR PFAM: PF00739; Gemini_AL1; 1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00228; GEMCOATL1.
FW NP-BIND 219 226 ATP (POTENTIAL).
SQ SEQUENCE 357 AA; 40678 MW; 939A568E1A63B2A7 CRC64;

Query Match 56.5%; Score 203; DB 1; Length 357;
Best Local Similarity 65.6%; Pred. No. 7.7e-16;
Matches 40; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 4 WGEFVDGRSGRGCGTSNDAAEAALNKSKEALQIREKIPAAALFOFINLNSLDRI 63
DB 111 RVSSOLDGRSANGGQGSNDAAVAEALNKSKEALNLRKAPKDYILFOFINLNSLDRI 170

QY 64 F 64
DB 171 F 171

RESULT 14
VAL1_SICV STANDARD; PRT: 347 AA.
AC P29048;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DE ALL PROTEIN
OS Squash leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; begomovirus.
RN [1]
EX MEDLINE:91024449; PubMed:1984669;
RA Lazarowitz S.G., Lazdins J.H.;
RT "Nucleotide sequence of the cloned genomic components of a bipartite squash leaf curl geminivirus with a broad
RT host range phenotype.";
RL Virology 180:58-69(1991).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: M38183; ARC32410.1; ALT_INIT7.

```

```

DR PIR: C36785; QOCVSL.
DR INTERPRO: IPR001191; -.
DR PFAM: PF00739; Gemini_AL1; 1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00228; GEMCOATL1.
KW ATP-binding; 218 225 ATP (POTENTIAL).
PT NP-BIND
SQ SEQUENCE 347 AA; 39110 MW; AFDAEBDE12110E CRC64;

Query Match 52.9%; Score 118; DB 1; Length 347;
Best Local Similarity 37.3%; Pred. No. 1.2e-06;
Matches 25; Conservative 12; Mismatches 25; Indels 4; Gaps 1;

QY 5 GEFVDGRSGRGCGTSNDAAEAALNKSKEALQIREKIPAAALFOFINLNSLDRI 64
DB 116 QTKYKVS-----GSKSKODYTHNVAAGSAGEALQIKAGDPKFTFYVYHLLANVERLF 171

QY 65 DKTPPEP 70
DB 172 QKPPPEP 177

RESULT 15
VEL_LHPV66 STANDARD; PRT: 630 AA.
ID VEL_LHPV66
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE REPLICATION PROTEIN E1.
DE REPLICATION PROTEIN E1.
GN E1.
OS Human papillomavirus type 66.
OC Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Papillomavirus.
RN [1]
RP SEQUENCE FROM N.A.
RA Delius H.;
RL Submitted (OCT-1995) to the EMBL/GenBank/ODS databases.
CC -1- FUNCTION: ATP-DEPENDENT DNA HELICASE REQUIRED FOR INITIATION OF
CC VIRAL DNA REPLICATION. IT FORMS A COMPLEX WITH THE VIRAL E2
CC PROTEIN. THE E1-E2 COMPLEX BINDS TO THE REPLICATION ORIGIN WHICH
CC CONTAINS SEVERAL DNA REPLICATION ORIGIN PROTEINS.
CC -1- SUBCELLULAR LOCATION: NUCLEUS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: U31794; AAA79501.1; -.
DR INTERPRO: IPR001177; -.
DR PFAM: PF00519; E1; 1.
DR PFAM: PF00524; E1_N; 1.
KW Early protein; DNA replication; Helicase; ATP-binding; DNA-binding;
FW NP-BIND 458 465 ATP (POTENTIAL).
PT NP-BIND
SQ SEQUENCE 630 AA; 72065 MW; 22DDDA5934F791B CRC64;

Query Match 18.1%; Score 65; DB 1; Length 630;
Best Local Similarity 25.5%; Pred. No. 6.8;
Matches 13; Conservative 14; Mismatches 24; Indels 0; Gaps 0;

QY 19 QFSNDAAEAALNKSKEALQIREKIPAAALFOFINLNSLDRI 69
DB 269 KTTKLSLSLNVQFQMLTOPPKIRSPALVYFKTKAMSNISFVGVGTPE 319

```

Search Completed: February 3, 2001, 02:24:27  
Job time: 639 sec

---

**THIS PAGE BLANK (USPTO)**

Genome version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: February 3, 2001, 02:17:38 ; Search time 118.74 seconds  
(without alignments)  
40.029 Million cell updates/sec

Title: US-09-289-346a-5

Best score: 159

Sequence: 1 TLVWGEFQVDSARGGCOT.....FQPHNLNLDIFDKTPEP 70

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR.66:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than the observed score of the best hit being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	343	95.5	352	1 QOCVLI	Al1 protein - toma
2	247	68.8	361	1 QOCVPT	Al1 protein - tola
3	243	67.7	358	2 S07594	hypothetical prote
4	232	64.6	362	1 JQ1887	Al1 protein - toma
5	226	63.0	359	2 S39211	gene C1 protein -
6	224	62.4	349	2 JQ2300	replicase - pepper
7	223	62.4	349	2 S31873	Al1 protein - pepp
8	219	61.9	349	2 S31873	Al1 protein - pep
9	216	60.2	351	2 JQ2327	Al1 protein - tola
10	216	60.2	358	1 JQ1870	Al1 protein - tola
11	215	59.9	355	1 QOCVW1	AVI protein - abut
12	213	59.3	359	2 S39235	gene C1 protein -
13	212	59.1	385	2 S28360	Al1 protein - beet
14	209	58.2	360	2 S59885	replication-associ
15	203	56.5	357	1 QOCVLI	Al1 protein - toma
16	203	56.5	357	1 QOCVLI	Al1 protein (clone
17	118	33.9	317	1 QOCVSL	Al1 protein - tola
18	65	18.1	597	2 JQ1419	FC gamma (19G) rec
19	63.5	17.7	1713	2 A55347	adhesive ligand ep
20	62	17.3	160	2 G82060	hypothetical prote
21	61	17.0	631	2 S30505	E1 protein - human
22	61	17.0	840	2 T30175	probable large ATP
23	60	16.9	423	2 A43394	GAP-binding regula
24	60.5	16.9	423	2 A43394	hypothetical prote
25	60.5	16.9	447	2 T12544	hypothetical prote
26	60.5	16.9	1070	1 A54600	1-phosphatidylinos
27	60	16.7	397	2 B71078	probable NADH oxid
28	59.5	16.6	706	2 H71078	hypothetical prote
29	59.5	16.6	1053	2 T05483	probable ubiquitin

## ALIGNMENTS

### RESULT 1

QOCVLI

Al1 protein - tomato golden mosaic virus

C:Species: tomato golden mosaic virus

A:Note: host Nicotiana sp. (tobacco)

C:Date: 28-Aug-1985 #sequence\_revision 28-Aug-1985 #text\_change 06-Apr-1994

C:Accession: A04170 Stejn, V.E.; Coutts, R.H.A.; Buck, K.W.

Hamilton, 2197-2205, 1984

EMBL J, 2197-2205, 1984

A:Title: Complete nucleotide sequence of the infectious cloned DNA components of toma

A:Reference number: A04163

A:Accession: A04170

A:Molecule type: DNA

A:Residues: 1-352 <HAMP>

C:Comment: The genome consists of two circular, single-stranded DNA components, DNA A

C:Genetics:

A:Map position: segment A

C:Superfamily: tomato golden mosaic virus Al1 protein

Query Match 95.5%; Score 343; DB 1; Length 352;  
Best Local Similarity 95.7%; Prev. No. 146-31;  
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 TLVWGEFQVDSARGGCOTSDNAAEALNASSKEALQIIRKIPAAALFQPHNLNSNL 60

Db 111 TLVWGEFQVDSARGGCOTSDNAAEALNASSKEALQIIRKIPKYLQPHNLNSNL 170

Oy 61 DRIFDKTPEP 70

Db 171 DRIFDKTPEP 180

### RESULT 2

QOCVPT

Al1 protein - potato yellow mosaic virus (isolate Venezuela)

C:Species: potato yellow mosaic virus

C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Jun-2000

C:Accession: J00364

C:Comments: R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.

J. Gen. Virol. 72, 1515-1520, 1991

A:Title: The nucleotide sequence of the infectious cloned DNA components of potato y

A:Reference number: J00362; MUID:9111403

A:Accession: J00364

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-351 <HAMP>

C:Cross-references: GB:D00940; NID:9222458; PIDN:BA00782.1; PID:9222459

C:Genetics:

A:Map position: segment A

C:Superfamily: tomato golden mosaic virus Al1 protein

```

Query Match          58.8%  Score 247;  Db 1;  Length 361;
Best Local Similarity 68.1%  Pred. No. 1.1e-18;
Matches 47;  Conservative 10;  Mismatches 12;  Indels 0;  Gaps 0;

QY  1 TLVWGEFVDGSRGCGCOTSNDAAALNASSKEALQIIREKIPAAALFQFHNLSNL 60
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db  110 TLVWGEFVDGSRGCGCOTVNDAAALNASSKEALQIIREKIPAAALFQFHNLSNL 169
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY  61 DRIFDKTPEP 69
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db  170 DRIFDKTPEP 178
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

RESULT 3
S07594
Synthetical protein, 40.4K - cassava latent virus (Nigerian isolate)
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C:Accession: S07594
R:Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.
Nucleic Acids Res. 18, 197-198, 1990
A:Title: Nucleotide sequence of the infectious cloned DNA components of African cassava
A:Reference number: S07594
A:Accession: S07594
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-358 <NOR>
A:Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376
C:Genetics:
A:Map position: segment DNAL
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match          67.7%  Score 243;  Db 2;  Length 358;
Best Local Similarity 64.3%  Pred. No. 3.1e-20;
Matches 45;  Conservative 11;  Mismatches 14;  Indels 0;  Gaps 0;

QY  1 TLVWGEFVDGSRGCGCOTSNDAAALNASSKEALQIIREKIPAAALFQFHNLSNL 60
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db  109 TLVWGEFVDGSRGCGCOTNDAAALNASSKEALQIIREKIPAAALFQFHNLSNL 168
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY  61 DRIFDKTPEP 70
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db  169 DRIFDKTPEP 178
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

RESULT 4
AL1897
AL1 protein - tomato yellow leaf curl virus (strain Australia)
N:Alternate names: CL protein
C:Species: tomato yellow leaf curl virus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
R:Dry, I.B.; Rigden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.
Mol. Cell. Biol. 14, 47-53, 1993
A:Title: Nucleotide sequence of the genome organization of tomato leaf curl geminivirus.
A:Reference number: JQ1887; MUID:9313978
A:Accession: JQ1887
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-362 <DR>
A:Cross-references: GB:S53251
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match          64.6%  Score 232;  Db 1;  Length 362;
Best Local Similarity 55.3%  Pred. No. 5.6e-19;
Matches 47;  Conservative 8;  Mismatches 14;  Indels 16;  Gaps 1;

QY  1 TLVWGEFVDGSRGCGCOTSNDAAALNASSKEALQIIREKIPAAALFQFHNLSNL 60
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

```

Db  110 TLVWGEFVDGSRGCGCOTSNDAAALNASSKEALQIIREKIPAAALFQFHNLSNL 169
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY  61 DRI-----FDKTPPE 69
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db  170 DRIFDKTPEP 178
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

RESULT 5
gene Cl protein - tomato yellow leaf curl virus
C:Species: tomato yellow leaf curl virus
C:Date: 06-Jan-1993 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
C:Accession: S39211
R:Morris, B.; Hidalgo, E.; Accotto, G.; Mortones, E.
Submitted to the EMBL Data Library, August 1993
A:Description: High similarity among the tomato yellow leaf curl virus isolates from
A:Reference number: S39209
A:Accession: S39211
A:Status: preliminary
A:Accession: S39211
A:Residues: 1-359 <NOR>
A:Cross-references: EMBL:225751; NID:g433655; PIDN:CAA81026.1; PID:g433658
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match          63.0%  Score 226;  Db 2;  Length 359;
Best Local Similarity 60.9%  Pred. No. 2.7e-18;
Matches 42;  Conservative 9;  Mismatches 18;  Indels 0;  Gaps 0;

QY  2 LWVGEFVDGSRGCGCOTSNDAAALNASSKEALQIIREKIPAAALFQFHNLSNL 61
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db  111 LWVGEFVDGSRGCGCOTNDAAALNASSKEALQIIREKIPAAALFQFHNLSNL 170
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY  62 DRIFDKTPEP 70
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db  171 DRIFDKTPEP 179
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

RESULT 6
JQ2300
replicase - pepper huasteco virus (component A)
N:Alternate names: ORF AL1 protein
C:Species: pepper huasteco virus
C:Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Sep-1999
C:Accession: JQ2300
R:Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante,
J. Gen. Virol. 74, 2225-2231, 1993
A:Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b
A:Reference number: JQ2299; MUID:94015007
A:Accession: JQ2300
A:Status: type strain
A:Residues: 1-349 <TOR>
A:Cross-references: GB:X70418; NID:g61023; PIDN:CAA9856.1; PID:g61025
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match          62.4%  Score 224;  Db 2;  Length 349;
Best Local Similarity 61.4%  Pred. No. 4.5e-18;
Matches 43;  Conservative 11;  Mismatches 16;  Indels 0;  Gaps 0;

QY  1 TLVWGEFVDGSRGCGCOTSNDAAALNASSKEALQIIREKIPAAALFQFHNLSNL 60
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db  110 TLVWGEFVDGSRGCGCOTSNDAAALNASSKEALQIIREKIPAAALFQFHNLSNL 169
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY  61 DRIFDKTPEP 70
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db  170 DRIFDKTPEP 179
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

RESULT 7
S31875
AL1 protein - pepper rizado amarillo virus
C:Species: pepper rizado amarillo virus

```





QY 1 TLWGEFVQDGRSARGGCGTSDNAAAEALNASKEEALQIIRKIPAAALFQFHNLNSML 60  
 Db 110 TLWGEFVQDGRSARGGCGTSDNAAAEALNASKEEALQIIRKIPAAALFQFHNLNSML 169  
 QY 61 DRIFDKTPEP 70  
 Db 170 ERIFAKAPEP 179

RESULT 12  
 S39235  
 gene Cl protein - tomato yellow leaf curl virus  
 C:Species: tomato yellow leaf curl virus  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 08-Sep-1997  
 C:Accession: S39235  
 R:Crespi, S.; Noris, E.; Valira, A.; Bosco, D.; Accotto, G.  
 submitted to the EMBL Data Library, December 1993  
 A:Reference number: S39235  
 A:Accession: S39235  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-359 <R>  
 A:Cross-references: EMBL:X04144; NID:g1041671; PID:g1334964  
 C:Superfamily: tomato golden mosaic virus ALL protein

Query Match 59.3%; Score 213; DB 2; Length 359;  
 Best Local Similarity 56.5%; Pred. No. 8.4e-17;  
 Matches 39; Conservative 12; Mismatches 18; Indels 0; Gaps 0;

QY 2 LVWGEFVQDGRSARGGCGTSDNAAAEALNASKEEALQIIRKIPAAALFQFHNLNSMLD 61  
 Db 111 LVWGEFVQDGRSARGGCGTSDNAAAEALNASKEEALQIIRKIPAAALFQFHNLNSMLD 170  
 QY 62 RIFDKTPEP 70  
 Db 171 KVFQVPPAP 179

RESULT 13  
 S28360  
 ALL protein - beet curly top virus  
 C:Species: beet curly top virus  
 C:Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 20-Sep-1999  
 C:Accession: S28360  
 R:Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.  
 submitted to the EMBL Data Library, December 1993  
 A:Reference number: S28360  
 A:Accession: S28360  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-385 <STA>  
 A:Cross-references: GB:X24597; EMBL:X04144; NID:g210678; PID:AAA42751.1; PID:g210679  
 C:Superfamily: tomato golden mosaic virus ALL protein

Query Match 59.1%; Score 212; DB 2; Length 385;  
 Best Local Similarity 55.7%; Pred. No. 1.2e-16;  
 Matches 39; Conservative 15; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLWGEFVQDGRSARGGCGTSDNAAAEALNASKEEALQIIRKIPAAALFQFHNLNSML 60  
 Db 137 TLWGEFVQDGRSARGGCGTSDNAAAEALNASKEEALQIIRKIPAAALFQFHNLNSML 196  
 QY 61 DRIFDKTPEP 70  
 Db 197 OKIFORPPDP 206

RESULT 14

S59885  
 replication-associated protein Cl - tomato yellow leaf curl virus  
 C:Species: tomato yellow leaf curl virus  
 C:Date: 14-Jun-1998 #sequence\_revision 01-Mar-1996 #text\_change 20-Sep-1999  
 C:Accession: S59885  
 R:Hong, Y.; Harrison, B.D.  
 submitted to the EMBL Data Library, February 1995  
 A:Description: Nucleotide sequences from tomato leaf curl viruses from different countries  
 d:geminiviruses.  
 A:Reference number: S58346  
 A:Accession: S59885  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-360 <R>  
 A:Cross-references: EMBL:X48182; NID:g944838; PIDN:CAA89229.1; PID:g974211  
 C:Superfamily: tomato golden mosaic virus ALL protein

Query Match 58.2%; Score 209; DB 2; Length 360;  
 Best Local Similarity 52.1%; Pred. No. 2.4e-16;  
 Matches 41; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

QY 4 WGEFVQDGRSARGGCGTSDNAAAEALNASKEEALQIIRKIPAAALFQFHNLNSMLDRI 63  
 Db 113 WGEFVQDGRSARGGCGTSDNAAAEALNASKEEALQIIRKIPAAALFQFHNLNSMLDRI 172  
 QY 64 FDKTPEP 69  
 Db 173 FTSAP 178

RESULT 15

QCVCL  
 ALL protein - tomato yellow leaf curl virus  
 C:Species: tomato yellow leaf curl virus  
 C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 23-Jul-1999  
 C:Accession: D40779  
 R:Navot, N.; Pichersky, E.; Zeldin, M.; Zamir, D.; Czosnek, H.  
 submitted to the EMBL Data Library, December 1993  
 A:Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin  
 virology 185, 151-161, 1991  
 A:Reference number: A40779; MUID:92024070  
 A:Accession: D40779  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-357 <NAV>  
 A:Cross-references: GB:X15656; NID:g62204; PIDN:CAA33688.1; PID:g62207  
 C:Superfamily: tomato golden mosaic virus ALL protein

Query Match 56.5%; Score 203; DB 1; Length 357;  
 Best Local Similarity 55.6%; Pred. No. 1.2e-15;  
 Matches 40; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 4 WGEFVQDGRSARGGCGTSDNAAAEALNASKEEALQIIRKIPAAALFQFHNLNSMLDRI 63  
 Db 111 WGEFVQDGRSARGGCGTSDNAAAEALNASKEEALQIIRKIPAAALFQFHNLNSMLDRI 170  
 QY 64 F 64  
 Db 171 F 171

Search completed: February 3, 2001, 02:17:39  
 Job time: 2211 sec



[illegible]

RC STRAIN-JAMAICA;  
 RA Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;  
 RL Plant Dis., 81:1251-1258(1997).  
 DR PFAM: PF00799; GEMIN1\_1; 1.  
 DR INTERPRO: IPRO01301; -.  
 DR PFAM: PF00799; GEMIN1\_1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATL1.  
 FT NON\_TER 185 185  
 SQ SEQUENCE 185 AA; 20975 MW; 3913650A02545EE1 CRC64;

Query Match 75.5%; Score 271; DB 12; Length 185;  
 Best Local Similarity 71.4%; Pred. No. 4.3e-24;  
 Matches 50; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

Qy 1 TLVNGFQVQDGRSARGCQTSNDAAAEALNASSKEEALQIIRKIPAAALFQFHINLSNL 60  
 Db 89 TLVNGFQVQDGRSARGCQTSNDAAAEALNASSKEEALQIIRKIPAAALFQFHINLSNL 148

Qy 61 DRIFDKTP 70  
 Db 149 DRIFSKPP 158

RESULT 6  
 ID 09Q555 PRELIMINARY; PRT; 364 AA.  
 AC 09Q555;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE REPLICATION INITIATION PROTEIN AC1.  
 OS sweet potato leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=100755;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;  
 RT Detection of a geminivirus infecting sweet potato in the United States  
 RI Plant Dis. 82:1253-1257(1998).  
 DR EMBL: AF104036; AAD47173.1; -.  
 DR INTERPRO: IPRO01191; -.  
 DR INTERPRO: IPRO01301; -.  
 DR PFAM: PF00799; GEMIN1\_1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATL1.  
 FT NON\_TER 364 364  
 SQ SEQUENCE 364 AA; 40680 MW; 5F79752431A09D6E CRC64;

Query Match 75.5%; Score 271; DB 12; Length 364;  
 Best Local Similarity 79.4%; Pred. No. 9.3e-24;  
 Matches 54; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

Qy 1 TLVNGFQVQDGRSARGCQTSNDAAAEALNASSKEEALQIIRKIPAAALFQFHINLSNL 60  
 Db 110 TLVNGFQVQDGRSARGCQTSNDAAAEALNASSKEEALQIIRKIPAAALFQFHINLSNL 169

Qy 61 DRIFDKTP 68  
 Db 170 DRIFSKPP 177

RESULT 7  
 ID P88975 PRELIMINARY; PRT; 149 AA.  
 AC P88975;  
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)

DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).  
 GN AC1.  
 OS Macrotidium golden mosaic geminivirus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=51676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAMAICA;  
 RA Roye M.E.;  
 RL Thesis (1996); Biochemistry, University of the West Indies, Jamaica.  
 DR EMBL: U75278; AAB56919.1; -.  
 DR EMBL: U75278; AAB56919.1; -.  
 DR INTERPRO: IPRO01301; -.  
 DR PFAM: PF00799; GEMIN1\_1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATL1.  
 FT NON\_TER 149 149  
 SQ SEQUENCE 149 AA; 16785 MW; E4CF5EED4C9CD508 CRC64;

Query Match 73.5%; Score 264; DB 12; Length 149;  
 Best Local Similarity 70.0%; Pred. No. 2.2e-23;  
 Matches 49; Conservative 11; Mismatches 10; Indels 0; Gaps 0;

Qy 1 TLVNGFQVQDGRSARGCQTSNDAAAEALNASSKEEALQIIRKIPAAALFQFHINLSNL 60  
 Db 52 TLVNGFQVQDGRSARGCQTSNDAAAEALNASSKEEALQIIRKIPAAALFQFHINLSNL 111

Qy 61 DRIFDKTP 70  
 Db 112 DRIFSKDP 121

RESULT 8  
 ID 09Z089 PRELIMINARY; PRT; 190 AA.  
 AC 09Z089;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE C1 AND C4 GENES, CLONE YOKOHAMA3-1,  
 DE PARTIAL AND COMPLETE CDS (FRAGMENT).  
 OS Tobacco leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=67762;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-YOKOHAMA3;  
 RA Ooi K., Oshita S., Ishii I., Yahara T.;  
 RT Geminivirus infecting tobacco in Japan.  
 RI Plant Res. 110:247-257(1997).  
 DR EMBL: AB001315; BAA34033.1; -.  
 DR INTERPRO: IPRO01191; -.  
 DR INTERPRO: IPRO01301; -.  
 DR PFAM: PF00799; GEMIN1\_1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATL1.  
 FT NON\_TER 190 190  
 SQ SEQUENCE 190 AA; 21432 MW; AAC093D1D1610PAD CNC64;

Query Match 73.5%; Score 264; DB 12; Length 190;  
 Best Local Similarity 62.4%; Pred. No. 2.9e-23;  
 Matches 53; Conservative 6; Mismatches 10; Indels 16; Gaps 1;

Qy 1 TLVNGFQVQDGRSARGCQTSNDAAAEALNASSKEEALQIIRKIPAAALFQFHINLSNL 60  
 Db 85 TLVNGFQVQDGRSARGCQTSNDAAAEALNASSKEEALQIIRKIPAAALFQFHINLSNL 144

```

QY 61 DRI-----FDKTP6 9
|||
Db 145 DRIFAPPLEVFCVFASFDQVPE 169
|||
|||

RESULT 9
ID Q92084 PRELIMINARY: PRT: 190 AA.
AC Q92084 (TEMBLrel. 10, Created)
DT 01-MAY-1999 (TEMBLrel. 10, Last sequence update)
DT 01-MAY-2000 (TEMBLrel. 10, Last annotation update)
DE C1 AND C4 GENES; CLONE YKOHAMA5-
DE PARTIAL AND COMPLETE CDS (FRAGMENT).
GN C1.
OS tobacco leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=67762;
RN 111_TaxID=67762;
RP SEQUENCE FROM N.A.
RC STRAIN-YKOHAMA5.
RA Ooi K., Ohshita S., Ishii I., Yahara T.;
RL "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
J. Plant Res. 110:247-257(1997).
DR EMBL: AB001318; BAA34039.1; -.
DR FRANK: PR001301; L:ALI; 1.
DR PFAM: PF00759; GEMCOATL1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00228; GEMCOATL1.
FT NON-TER 1
FT NON-TER 190
SQ SEQUENCE 190 AA; 21444 MW; AACIC2943E3F01AD CRC64;

Query Match 73.5%; Score 264; DB 12; Length 190;
Best Local Similarity 62.4%; Pred. No. 2,9e-23;
Matches 53; Conservative 6; Mismatches 10; Indels 16; Gaps 1;

QY 1 TLWGEFQVDSARGCGQTSNDAAALNAASKKEALQIREKIPAAALFQHNLSNL 60
|||
Db 85 TLEWGTQVDSARGCGQTSNDAAALNAASKKEALQIREKLPKDFIFQHNLSNL 144
|||
|||

QY 61 DRI-----FDKTP6 9
|||
Db 145 DRIFAPPLEVFCVFASFDQVPE 169
|||
|||

RESULT 10
ID Q9YLAA PRELIMINARY: PRT: 233 AA.
AC Q9YLAA (TEMBLrel. 10, Created)
DT 01-MAY-1999 (TEMBLrel. 10, Last sequence update)
DT 01-MAY-2000 (TEMBLrel. 10, Last annotation update)
DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN REP. proteinium golden mosaic geminivirus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=51676;
RN 111_TaxID=51676;
RP SEQUENCE FROM N.A.
RC STRAIN-JAMAICA STRAIN 1;
RA Roye M.E.;
RT Genetic diversity and phylogeny of whitefly-transmitted geminiviruses
from Jamaica.
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN-JAMAICA STRAIN 1;
RA Roye M.E., McLaughlin W.A., Maxwell D.P.;
RT Molecular characterization of two distinct geminiviruses infecting M.
ethiopioides from Jamaica.;

```

```

RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF098940; AAD17850.1; -.
DR INTERPRO: IPR001191; -.
DR TRN: TRN001301; L:ALI; 1.
DR PFAM: PF00759; GEMCOATL1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00228; GEMCOATL1.
FT NON-TER 233
SQ SEQUENCE 233 AA; 26355 MW; AA490AF4D216GA02 CRC64;

Query Match 73.5%; Score 264; DB 12; Length 233;
Best Local Similarity 70.0%; Pred. No. 3,7e-23;
Matches 49; Conservative 11; Mismatches 10; Indels 0; Gaps 0;

QY 1 TLWGEFQVDSARGCGQTSNDAAALNAASKKEALQIREKIPAAALFQHNLSNL 60
|||
Db 110 TLEWGTQVDSARGCGQTSNDAAALNAASKKEALQIREKLPKDFIFQHNLSNL 169
|||
|||

QY 61 DRI-----FDKTP6 9
|||
Db 170 DRIFAPPLEVFCVFASFDQVPE 169
|||
|||

RESULT 11
Q9M827 PRELIMINARY: PRT: 190 AA.
AC Q9M827 (TEMBLrel. 12, Created)
DT 01-NOV-1999 (TEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TEMBLrel. 12, Last sequence update)
DT 01-MAY-2000 (TEMBLrel. 13, Last annotation update)
DE C1 PROTEIN (FRAGMENT).
GN C1.
OS tobacco leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=67762;
RN 111_TaxID=67762;
RP SEQUENCE FROM N.A.
RC STRAIN-GORAL;
RA Ooi K., Ohshita S., Ishii I., Yahara T.;
RT "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
J. Plant Res. 110:247-257(1997).
DR EMBL: AB001318; BAA34039.1; -.
DR FRANK: PR001301; L:ALI; 1.
DR INTERPRO: IPR001191; -.
DR PFAM: PF00759; GEMCOATL1.
DR PFAM: PF00799; GEMCOATL1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00228; GEMCOATL1.
FT NON-TER 1
FT NON-TER 190
SQ SEQUENCE 190 AA; 21444 MW; 93C3742A8EBDB7BE CRC64;

Query Match 73.3%; Score 263; DB 12; Length 190;
Best Local Similarity 62.4%; Pred. No. 3,8e-23;
Matches 53; Conservative 6; Mismatches 10; Indels 16; Gaps 1;

QY 1 TLWGEFQVDSARGCGQTSNDAAALNAASKKEALQIREKIPAAALFQHNLSNL 60
|||
Db 85 TLEWGTQVDSARGCGQTSNDAAALNAASKKEALQIREKLPKDFIFQHNLSNL 144
|||
|||

QY 61 DRI-----FDKTP6 9
|||
Db 145 DRIFAPPLEVFCVFASFDQVPE 169
|||
|||

RESULT 12
ID Q9Z0C4 PRELIMINARY: PRT: 208 AA.
AC Q9Z0C4 (TEMBLrel. 10, Created)
DT 01-MAY-1999 (TEMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TEMBLrel. 10, Last sequence update)
DT 01-MAY-2000 (TEMBLrel. 13, Last annotation update)

```

C1 AND C4 GENES, CLONE ABURA3-1, PARTIAL  
DE AND COMPLETE CDS (FRAGMENT).  
GN C1.  
OS tobacco leaf curl virus.  
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
ON NCBI\_TaxID=67762;  
OX [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ABURA3;  
RA Ooi K., Ohshita S., Ishii I., Yahara T.;  
RT Molecular phylogeny of geminivirus infecting wild plants in Japan.\*;  
RL Plant Res. 110:247-257(1997).  
RM EMBL: AB001296, BAA33996.1; -;  
DR INTERPRO: IPRO01191; -;  
DR INTERPRO: IPRO01191; -;  
DR PFAM: PF007599; Gemin1\_AL1; 1.  
DR PRINTS: PR00227; GEMCOATL1.  
DR PRINTS: PR00228; GEMCOATL1.  
DR NON\_TER 1 208  
FT NON\_TER 208  
SQ SEQUENCE 208 AA; 23526 MW; 249CC1D8729C72D CRC64;

Query Match 73.38; Score 263; DB 12; Length 208;  
Best Local Similarity 62.48; Pred. No. 4, 28-23;  
Matches 53; Conservative 6; Mismatches 10; Indels 16; Gaps 1;  
Qy 1 TLWGFQVQDGRSARGGCGTNDAAAEALNASSKEEALQIIRKIPAAALFQFHNLNSL 60  
Db 102 TLENGTFTQDGRSARGGCONACAEALNASSKADALAIIRKLPKDFIQYHNLNSL 161  
Qy 61 DRI-----FKTPE 69  
Db 162 DRIFAPLEVPVCPSSSFDQVPE 186

RESULT 13  
Q920C0  
ID Q920C0 PRELIMINARY; PRT: 208 AA.  
AC Q920C0;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
DE C1 AND C4 GENES, CLONE ABURA3-1, PARTIAL AND COMPLETE CDS (FRAGMENT).  
GN C1.  
OS tobacco leaf curl virus.  
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
ON NCBI\_TaxID=67762;  
OX [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ABURA3;  
RA Ooi K., Ohshita S., Ishii I., Yahara T.;  
RT Molecular phylogeny of geminivirus infecting wild plants in Japan.\*;  
RL Plant Res. 110:247-257(1997).  
RM EMBL: AB001296; BAA33996.1; -;  
DR INTERPRO: IPRO01191; -;  
DR INTERPRO: IPRO01191; -;  
DR PFAM: PF007599; Gemin1\_AL1; 1.  
DR PRINTS: PR00227; GEMCOATL1.  
DR PRINTS: PR00228; GEMCOATL1.  
DR NON\_TER 1 208  
FT NON\_TER 208  
SQ SEQUENCE 208 AA; 23486 MW; E301135F799C3DAD CRC64;

Query Match 72.78; Score 261; DB 12; Length 208;  
Best Local Similarity 61.28; Pred. No. 7, 3e-23;  
Matches 52; Conservative 7; Mismatches 10; Indels 16; Gaps 1;  
Qy 1 TLWGFQVQDGRSARGGCGTNDAAAEALNASSKEEALQIIRKIPAAALFQFHNLNSL 60  
Db 102 TLENGTFTQDGRSARGGCONACAEALNASSKADALAIIRKLPKDFIQYHNLNSL 161

Qy 61 DRI-----FKTPE 69  
Db 162 DRIFAPLEVPVCPSSSFDQVPE 186  
RESULT 14  
Q920B8  
ID Q920B8 PRELIMINARY; PRT: 208 AA.  
AC Q920B8;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
DE C1 AND C4 GENES, CLONE AMG-2 (B15A), PARTIAL AND COMPLETE CDS (FRAGMENT).  
GN C1.  
OS tobacco leaf curl virus.  
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
ON NCBI\_TaxID=67762;  
OX [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AMG15A;  
RA Ooi K., Ohshita S., Ishii I., Yahara T.;  
RT Molecular phylogeny of geminivirus infecting wild plants in Japan.\*;  
RL Plant Res. 110:247-257(1997).  
RM EMBL: AR001297; BAA33998.1; -;  
DR INTERPRO: IPRO01191; -;  
DR INTERPRO: IPRO01191; -;  
DR PFAM: PF007599; Gemin1\_AL1; 1.  
DR PRINTS: PR00227; GEMCOATL1.  
DR PRINTS: PR00228; GEMCOATL1.  
DR NON\_TER 1 208  
FT NON\_TER 208  
SQ SEQUENCE 208 AA; 23472 MW; 629D0DEF7C9956AA CRC64;

Query Match 72.78; Score 261; DB 12; Length 208;  
Best Local Similarity 61.28; Pred. No. 7, 3e-23;  
Matches 52; Conservative 7; Mismatches 10; Indels 16; Gaps 1;  
Qy 1 TLWGFQVQDGRSARGGCGTNDAAAEALNASSKEEALQIIRKIPAAALFQFHNLNSL 60  
Db 102 TLENGTFTQDGRSARGGCONACAEALNASSKADALAIIRKLPKDFIQYHNLNSL 161  
Qy 61 DRI-----FKTPE 69  
Db 162 DRIFAPLEVPVCPSSSFDQVPE 186

RESULT 15  
Q920B6  
ID Q920B6 PRELIMINARY; PRT: 208 AA.  
AC Q920B6;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
DE C1 AND C4 GENES, CLONE AMG-3 (A055), PARTIAL AND COMPLETE CDS (FRAGMENT).  
GN C1.  
OS tobacco leaf curl virus.  
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
ON NCBI\_TaxID=67762;  
OX [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AMGA055;  
RA Ooi K., Ohshita S., Ishii I., Yahara T.;  
RT Molecular phylogeny of geminivirus infecting wild plants in Japan.\*;  
RL Plant Res. 110:247-257(1997).  
RM EMBL: AB001296; BAA34000.1; -;  
DR INTERPRO: IPRO01191; -;  
DR INTERPRO: IPRO01191; -;  
DR PFAM: PF007599; Gemin1\_AL1; 1.  
DR PRINTS: PR00227; GEMCOATL1.

DR PRINTS: PR00228; GEMCOATCLVLL.  
FT NON\_TER 208 1  
FT NON\_TER 208 208  
SQ SEQUENCE 208 AA; 23586 MW; A5E61180BF05B9BE CNC64;

Query Match 72.1%; Score 259; DB 12; Length 208;  
Best Local Similarity 60.0%; Pred. No. 1.2e-22;  
Matches 51; Conservative 8; Mismatches 10; Indels 16; Gaps 1;

QY 1 TLVNGFVDGSRGCGQTSNDRAALNASSKEKALQIPEKIPAAALFQFHNLNSML 60  
Db 102 TLEKGTFTDHSRSGGCONADACALNNTAELSLITERLPGDFIQENLNSML 161

QY 61 DRI-----FDKTPK 69  
Db 162 DRIFSPPLEVFCPPSSSFQVPE 186

Search completed: February 3, 2001, 02:22:51  
Job time: 1888 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:22:51 ; Search time 180.33 Seconds  
(without alignments)  
45.497 Million cell updates/sec

Title: US-09-289-346a-6

Percent score: 357

Sequence: 1 TLVWGFQVDSRGSGCQT.....FAPALNSILDRIFDTPPEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPRMBL15.\*
- 1: sp.archaea.\*
  - 2: sp.bacteria.\*
  - 3: sp.fungi.\*
  - 4: sp.human.\*
  - 5: sp.invertebrate.\*
  - 6: sp.mammal.\*
  - 7: sp.mus.\*
  - 8: sp.organella.\*
  - 9: sp.phage.\*
  - 10: sp.plant.\*
  - 11: sp.todent.\*
  - 12: sp.virus.\*
  - 13: sp.vertibrate.\*
  - 14: sp.unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Score	Query Match	Length	DB ID	Description
1	272	76.2	226	12	009727 leonurus mo
2	272	76.2	226	12	09WHF6 tomato mild
3	270	75.6	361	12	067574 bean golden
4	268	75.1	225	12	090DB1 cowpea gold
5	268	75.1	364	12	090B55 sweet potat
6	267	74.8	225	12	090B55 sweet potat
7	257	72.0	119	12	090B55 sweet potat
8	257	72.0	233	12	09YL44 macroptiliu
9	251	70.3	234	12	039180 potato yell
10	248	69.5	190	12	092089 tobacco lea
11	248	69.5	190	12	092084 tobacco lea
12	247	69.2	190	12	09WH27 tobacco lea
13	245	68.2	208	12	0920C4 tobacco lea
14	245	68.2	208	12	0920C4 tobacco lea
15	245	68.0	208	12	0920B8 tobacco lea
16	243	68.1	208	12	0920B6 tobacco lea
17	242	67.8	203	12	0920B3 tobacco lea
18	242	67.8	349	12	088888 tomato pseu
19	242	67.8	363	12	073577 cotton leaf

0920a0	tobacco lea	20	239	66.9	208	12	0920A0
0920c6	tobacco lea	21	238	66.7	208	12	0920C6
072705	cotton leaf	22	235	65.8	363	12	072705
072719	cotton leaf	23	235	65.8	363	12	072719
072719	cotton leaf	24	235	65.8	363	12	072719
056816	chayote mos	25	233	65.3	362	12	056816
09yzv4	tomato yell	26	232	65.0	359	12	09YZV4
09yzv2	tomato yell	27	232	65.0	359	12	09YZV2
09yuk7	tomato yell	28	232	65.0	359	12	09YUX7
072692	beet curly	29	232	65.0	359	12	09YL27
072692	beet curly	30	231	64.7	353	12	072692
085942	beet curly	31	231	64.7	353	12	085942
085942	beet curly	32	231	64.7	353	12	085942
0920a7	tobacco lea	33	230	64.4	190	12	0920A7
091n48	okra enatio	34	230	64.4	362	12	091N48
091et7	cotton leaf	35	229	64.1	307	12	091ET7
072723	cotton leaf	36	229	64.1	361	12	072723
096620	african tom	37	228	63.9	231	12	096620
091n42	south afric	38	228	63.9	354	12	091N42
063419	beet curly	39	227	63.6	358	12	063419
091n48	okra enatio	40	227	63.6	358	12	091N48
09YU77	alpha rose	41	223	62.5	362	12	09YU77
074394	okra yellow	42	223	62.5	363	12	073494
091et1	cotton leaf	43	221	61.9	306	12	091ET1
09wr17	african cas	44	221	61.9	358	12	09WR17
09Jea9	cassava gem	45	221	61.9	358	12	09JFA9

#### ALIGNMENTS

RESULT 1  
ID 009727 PRELIMINARY; PRY: 226 AA.  
AC 009727: 1957 (T-EMBLrel. 04, Created)  
DT 01-JUL-1997 (T-EMBLrel. 04, Last sequence update)  
DT 01-OCT-2000 (T-EMBLrel. 15, Last annotation update)  
DE REP PROTEIN (FRAGMENT).  
GN REP.  
OS Leonurus mesasi virus.  
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=58177;  
RN 1  
RP STRAIN=LEW- BRAZIL 1;  
RC Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
RL Submitter J.C., Maxwell D.P.:  
FM EMBL: 092532; AAB51157.1; -;  
DR INTERPRO: IPR001191;  
DR INTERPRO: IPR001301; -; 1;  
DR SWISS-PROT: P00005; GEMINIVIRUS;  
DR PRINCS: PR00227; GEMINIVIRUS;  
DR PRINCS: PR00228; GEMINIVIRUS;  
FT NON\_TER 226  
SQ SEQUENCE 226 AA; 25617 MW; 73CDB6E766083FC5 CRC64;

Query Match  
Local similarity 76.2%; Score 272; DB 12; Length 226;  
Matches 54; Conservative 4; Mismatches 12; Indels 0;  
Gaps 0;

QY 1 TLVWGFQVDSRGSGCQTNSDAAEALNASKAEALQIIRKIPKYLFAFALNSNL 60  
|: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 111 TVENGFPQVDSRGSGGQTVNDAAEALNAPDKRTALQIIRKIPKYLFAFALNSNL 170  
QY 61 DRIFDTPPEP 70  
Db 171 DRIFNAKEP 180

RESULT 2  
Q9WHF6

```

ID OSWHF6 PRELIMINARY: PRT: 226 AA.
AC OSWHF6
DT 01-NOV-1999 (TEMBLrel. 12, Created)
DT 01-NOV-1999 (TEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)
DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN REP.
OS tomato mild mottle geminivirus.
OC Viruses: ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.
ON NCBI_TaxID=92943;
RN [1]
RC SEQUENCE FROM N.A.
RC SUBMITTER=H66-HK9.
RA Nakhla M.K., Nejla L., Ramirez P., Karkashian J.P., Doyle M.M.,
RA Maxwell D.P.;
RT "Molecular characterization and DNA-based detection methods for
RT vegetable-infecting geminiviruses in Central America.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF131071; AD33471.1; -.
DR INTERPRO: IPR001130; -.
DR PFAM: PF00759; Gemin_All: 1.
DR PRINTS: PR00227; GEMCOATALL.
DR PRINTS: PR00228; GEMCOATCLVL1.
DT NON_TER 226 226
SQ SEQUENCE 226 AA: 25941 MW; 2EA4116712B71A23 CRC64;

Query Match
Best Local Similarity 72.5%; Score 272; DB 12; Length 226;
Matches 51; Conservative 12; Mismatches 7; Indels 0; Gaps 0;

QY 1 TLVGFQVDSARGGCCTSDAAAEALMASKEEAQIIRKIPKYLFAFALNSNLI 60
Db 111 TLMGFQVDSARGGCCTSDAAAEALMASKEEAQIIRKIPKYLFAFALNSNLI 170
QY 61 DRIFKATPEP 70
Db 171 DRIFAKAPEP 180

RESULT 3
ID OS7574 PRELIMINARY: PRT: 361 AA.
AC OS7574;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-JUN-2000 (TEMBLrel. 14, Last annotation update)
DE PUTATIVE REPLICATIVE PROTEIN.
GN ALL.
OS can golden mosaic virus.
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.
ON NCBI_TaxID=10839;
RN [1]
RC SEQUENCE FROM N.A.
RA Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,
RA Maxwell D.P., Russell D.R.;
RL Phytopathology 81:980-985(1991).
RC SEQUENCE FROM N.A.
RA Gilbertson R.L., Hidayat S.R., Martinez R.T., Leong S.A., Faria J.C.,
RA Morales F.J., Maxwell D.P.;
RL Plant Dis. 75:336-342(1991).
RN [3]
RC SEQUENCE FROM N.A.
RA Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;
RN EMBL: M86865; AA446312.1; -.
DR INTERPRO: IPR001191; -.
DR PFAM: PF00799; Gemin_All: 1.
DR PRINTS: PR00227; GEMCOATALL.
DR PRINTS: PR00228; GEMCOATCLVL1.

```

```

DR PRODOM: PD000736; -. 1.
SQ SEQUENCE 361 AA: 41041 MW; 0094C7ACAF06B7B8 CRC64;

Query Match
Best Local Similarity 75.5%; Score 270; DB 12; Length 361;
Matches 51; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

QY 4 MGEFQVDSARGGCCTSDAAAEALMASKEEAQIIRKIPKYLFAFALNSNLI 63
Db 113 MGEFQVDSARGGCCTSDAAAEALMASKEEAQIIRKIPKYLFAFALNSNLI 172
QY 64 FKATPEP 70
Db 173 FKAPDP 179

RESULT 4
ID OSQDS5 PRELIMINARY: PRT: 225 AA.
AC OSQDS5;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TEMBLrel. 14, Last annotation update)
DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN REP.
OS can golden mosaic geminivirus.
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.
ON NCBI_TaxID=45203;
RN [1]
RC STRAIN=GMV-BR.
RA Faria J.C.;
RT "Partial nucleotide sequence of cowpea golden mosaic geminivirus from
RT Brazil.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF131071; AD33471.1; -.
DR INTERPRO: IPR001130; -.
DR PFAM: PF00799; Gemin_All: 1.
DR PRINTS: PR00227; GEMCOATALL.
DT NON_TER 225 225
SQ SEQUENCE 225 AA: 25766 MW; 109C6B6D8D15B5D0 CRC64;

Query Match
Best Local Similarity 75.1%; Score 268; DB 12; Length 225;
Matches 51; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 4 MGEFQVDSARGGCCTSDAAAEALMASKEEAQIIRKIPKYLFAFALNSNLI 63
Db 113 MGEFQVDSARGGCCTSDAAAEALMASKEEAQIIRKIPKYLFAFALNSNLI 172
QY 64 FKATPEP 70
Db 173 FKAPDP 179

RESULT 5
ID OSQDS5 PRELIMINARY: PRT: 364 AA.
AC OSQDS5;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TEMBLrel. 15, Last annotation update)
DE REPLICATION INITIATION PROTEIN AC1.
GN REP.
OS sweet potato leaf curl virus.
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.
ON NCBI_TaxID=100755;
RN [1]
RC SEQUENCE FROM N.A.

```

RA Iotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.:  
 RT "Detection of a geminivirus infecting sweet potato in the United  
 States"; 82-1253-1257 (1998).  
 DR EMBL AF104036:RAD04717.3.1; -  
 DR EMBL AF104036:RAD04717.3.1; -  
 DR INTERPRO: IPR001191; -  
 DR INTERPRO: IPR001191; -  
 DR PFAM: PF00799; Gemini\_1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATLVL1.  
 DR PRINTS: PR00228; GEMCOATLVL1.  
 SO SEQUENCE 364 AA; 40650 MW; 5F79752431A09D6E CRC64;

Query Match 75.1%; Score 268; DB 12; Length 364;  
 Best Local Similarity 79.4%; Pred. No. 3.3e-22;  
 Matches 54; Conservative 3; Mismatches 11; Indels 0; Gaps 0;  
 Oy 1 TLVNGFEVQDSARGCCOTSNDAARALNASSKEALQIIRKIPKYLFAFALNSNL 60  
 Db 110 TLVNGFEVQDSARGCCOTSNDAARALNASSKEALQIIRKIPKYLFAFALNSNL 169  
 Oy 61 DRIFDKTP 68  
 Db 170 DRIFDKTP 177

RESULT 6  
 ID Q98693 PRELIMINARY; PRT; 185 AA.  
 AC Q98693;  
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE REP PROTEIN (FRAGMENT).  
 OS Stida golden mosaic virus  
 GN REP.  
 OC Viruses; ssDNA viruses, Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=51034;  
 RN [1]  
 RC STRAIN-JAMAICA;  
 RA Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.:  
 DR EMBL AF098940:ADJ17650.1; -  
 DR EMBL AF098940:ADJ17650.1; -  
 DR INTERPRO: IPR001191; -  
 DR INTERPRO: IPR001191; -  
 DR PFAM: PF00799; Gemini\_1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATLVL1.  
 DR NON\_TER 185 185  
 FT NON\_TER 185 185  
 SO SEQUENCE 185 AA; 20975 MW; 3913850A025A5EE1 CRC64;

Query Match 74.2%; Score 265; DB 12; Length 185;  
 Best Local Similarity 70.0%; Pred. No. 3.2e-22;  
 Matches 49; Conservative 12; Mismatches 9; Indels 0; Gaps 0;  
 Oy 1 TLVNGFEVQDSARGCCOTSNDAARALNASSKEALQIIRKIPKYLFAFALNSNL 60  
 Db 89 TLVNGFEVQDSARGCCOTSNDAARALNASSKEALQIIRKIPKYLFAFALNSNL 148  
 Oy 61 DRIFDKTP 70  
 Db 149 DRIFDKTP 158

RESULT 7  
 ID P88975 PRELIMINARY; PRT; 149 AA.  
 AC P88975;  
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)

DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).  
 GN ACL.  
 OS Macrotellium golden mosaic geminivirus.  
 OC Viruses; ssDNA viruses, Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=51676;  
 RN [1]  
 RC STRAIN-JAMAICA;  
 RA Roye M.E.,  
 RL Thesis (1996), Biochemistry, University of the West Indies, Jamaica.  
 DR EMBL U75278:G00130.1; -  
 DR EMBL U75278:G00130.1; -  
 DR INTERPRO: IPR001191; -  
 DR PFAM: PF00799; Gemini\_1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATLVL1.  
 DR NON\_TER 149 149  
 FT NON\_TER 149 149  
 SO SEQUENCE 149 AA; 16785 MW; F4CF5ED4C9CD508 CRC64;

Query Match 72.0%; Score 257; DB 12; Length 149;  
 Best Local Similarity 68.6%; Pred. No. 2e-21;  
 Matches 48; Conservative 12; Mismatches 10; Indels 0; Gaps 0;  
 Oy 1 TLVNGFEVQDSARGCCOTSNDAARALNASSKEALQIIRKIPKYLFAFALNSNL 60  
 Db 52 TLVNGFEVQDSARGCCOTSNDAARALNASSKEALQIIRKIPKYLFAFALNSNL 111  
 Oy 61 DRIFDKTP 70  
 Db 112 DRIFDKTP 121

RESULT 8  
 ID Q9YL44 PRELIMINARY; PRT; 233 AA.  
 AC Q9YL44;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
 DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).  
 GN REP.  
 OS Macrotellium golden mosaic geminivirus  
 OC Viruses; ssDNA viruses, Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=51676;  
 RN [1]  
 RC STRAIN-JAMAICA STRAIN 1;  
 RA Roye M.E.,  
 RL Genetic diversity and phylogeny of whitefly-transmitted geminiviruses  
 DT Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RC STRAIN-JAMAICA STRAIN 1;  
 RC STRAIN-JAMAICA STRAIN 1;  
 RA Roye M.E., McLaughlin W.A., Maxwell D.P.:  
 RT Molecular characterization of two distinct geminiviruses infecting M.  
 fathyioides from Jamaica; The  
 DR EMBL AF098940:ADJ17650.1; -  
 DR EMBL AF098940:ADJ17650.1; -  
 DR INTERPRO: IPR001191; -  
 DR INTERPRO: IPR001191; -  
 DR PFAM: PF00799; Gemini\_1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATLVL1.  
 DR NON\_TER 233 233  
 FT NON\_TER 233 233  
 SO SEQUENCE 233 AA; 26355 MW; AA490AF4D2166A02 CRC64;

Query Match 72.0%; Score 257; DB 12; Length 233;  
 Best Local Similarity 68.6%; Pred. No. 3.3e-21;

Matches 48; Conservative 12; Mismatches 10; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDSARGGQTSNDAAALNASSKEALQIREKIPERYLFAFALNSNL 60  
 DB 110 TLVWGFQIDGSRGCGQTSNDAAALNASSKEALQIREKIPERYLFAFALNSNL 169

QY 61 DRIFDKTPEP 70  
 DB 170 DRIFDKMDEP 179

RESULT 9  
 ID Q93180 PRELIMINARY: PRT: 234 AA.  
 AC Q93180;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE REPLICATED-ASSOCIATED PROTEIN (FRAGMENT).  
 OS UL-1 AND C4 GENES, CLONE TOKUMAS-1,  
 OC VIRUSES; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10827;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=70M030 STRAIN;  
 RA Guzman F., Arredondo C.R., Emmatt D., Portillo R.J., Gilbertson R.L.;  
 DR EMBL: A0036553; BAA3823.1; (1997).  
 DE PARTIAL AND COMPLETE CDS (FRAGMENT).  
 DR INTERPRO: IPR001191; -.  
 DR INTERPRO: IPR001191; -.  
 DR PFAM: PF00799; Gemini\_All; 1.  
 DR PRINTS: PR00227; GEMCOATCLVL.  
 DR PRINTS: PR00228; GEMCOATCLVL.  
 FT NON\_TER 234 234  
 SQ SEQUENCE 234 AA; 26486 MW; 9ED8P0697105CD19 CRC64;

Query Match  
 Best Local Similarity 70.3%; Score 251; DB 12; Length 234;  
 Matches 47; Conservative 13; Mismatches 10; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDSARGGQTSNDAAALNASSKEALQIREKIPERYLFAFALNSNL 60  
 DB 110 TLVWGFQIDGSRGCGQTSNDAAALNASSKEALQIREKIPERYLFAFALNSNL 169

QY 61 DRIFDKTPEP 70  
 DB 170 DRIFDKMDEP 179

RESULT 10  
 ID Q92089 PRELIMINARY: PRT: 190 AA.  
 AC Q92089;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE UL-1 AND C4 GENES, CLONE TOKUMAS-1,  
 DE PARTIAL AND COMPLETE CDS (FRAGMENT).  
 GN C1  
 OS tobacco leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=67762;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=70K030 STRAIN;  
 RA Ochiai S., Ishii I., Yahara T.;  
 RT "Molecular phylogeny of geminivirus infecting wild plants in Japan.";  
 RL J. Plant Res. 110:247-257(1997).  
 DR EMBL: AB001315; BAA34033.1; -.  
 DR INTERPRO: IPR001191; -.  
 DR INTERPRO: IPR001191; -.  
 DR PFAM: PF00799; Gemini\_All; 1.

DR PRINTS: PR00227; GEMCOATCLVL.  
 DR PRINTS: PR00228; GEMCOATCLVL.  
 FT NON\_TER 190 190  
 SQ SEQUENCE 190 AA; 21432 MW; AAC093D01D1610PAD CRC64;

Query Match  
 Best Local Similarity 58.8%; Score 248; DB 12; Length 190;  
 Matches 50; Conservative 8; Mismatches 11; Indels 16; Gaps 1;

QY 1 TLVWGEFQVDSARGGQTSNDAAALNASSKEALQIREKIPERYLFAFALNSNL 60  
 DB 85 TLVWGFQIDGSRGCGQTSNDAAALNASSKEALQIREKIPERYLFAFALNSNL 144

QY 61 DRI-----FDKTP 69  
 DB 145 DRIFAPLEVFVCFPTASSEDQVPE 169

RESULT 11  
 ID Q92084 PRELIMINARY: PRT: 190 AA.  
 AC Q92084;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE UL-1 AND C4 GENES, CLONE TOKUMAS-1,  
 DE PARTIAL AND COMPLETE CDS (FRAGMENT).  
 GN C1  
 OS tobacco leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=67762;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=70K030 STRAIN;  
 RA Ochiai S., Ishii I., Yahara T.;  
 RT "Molecular phylogeny of geminivirus infecting wild plants in Japan.";  
 RL J. Plant Res. 110:247-257(1997).  
 DR EMBL: AB001318; BAA34039.1; -.  
 DR INTERPRO: IPR001191; -.  
 DR INTERPRO: IPR001191; -.  
 DR PFAM: PF00799; Gemini\_All; 1.  
 DR PRINTS: PR00227; GEMCOATCLVL.  
 DR PRINTS: PR00228; GEMCOATCLVL.  
 FT NON\_TER 190 190  
 SQ SEQUENCE 190 AA; 21444 MW; AAC1C2943E3F01AD CRC64;

Query Match  
 Best Local Similarity 59.5%; Score 248; DB 12; Length 190;  
 Matches 50; Conservative 8; Mismatches 11; Indels 16; Gaps 1;

QY 1 TLVWGEFQVDSARGGQTSNDAAALNASSKEALQIREKIPERYLFAFALNSNL 60  
 DB 85 TLVWGFQIDGSRGCGQTSNDAAALNASSKEALQIREKIPERYLFAFALNSNL 144

QY 61 DRI-----FDKTP 69  
 DB 145 DRIFAPLEVFVCFPTASSEDQVPE 169

RESULT 12  
 ID Q9W827 PRELIMINARY: PRT: 190 AA.  
 AC Q9W827;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE C1 PROTEIN (FRAGMENT).  
 GN C1  
 OS tobacco leaf curl virus.

Viruses: ssDNA viruses: Geminiviridae; Begomovirus.

```

OC NCBI_TaxID=67762;
RN [1]
RC SEQUENCE FROM N.A.
RA Ooi K., Ohshita S., Ishii I., Yahara T.;
RL J. Plant Res. 110:247-257(1997).
RT EMBL: AB001303; BAA34010.1; -.
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DE INTERPRO: IPRO01191; -.
DE C1 AND C4 GENES, CLONE ANG-1(B152),
DE PARTIAL AND COMPLETE CDS (FRAGMENT).
GN Clacaco leaf curl virus
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.
DR PRINTS: PR00228; GEMCOATCLVL1.
FT NON_TER 1 190
FT NON_TER 1 190
SQ SEQUENCE 190 AA; 21444 MW; 93C3742A8EBDB7EB CRC64;

```

```

Query Match 69.2%; Score 247; DB 12; Length 190;
Best Local Similarity 58.8%; Pred. No. 3.9e-20;
Matches 50; Conservative 8; Mismatches 11; Indels 16; Gaps 1;

QY 1 TLVGCFFVQDGRSNGCGTSDNAAAEALNASSKEEALQIIRKIPKYLFAFALNSNL 60
DB 85 TLEMGTFVQDGRSNGCGTSDNAAAEALNASSKEEALQIIRKIPKYLFAFALNSNL 144
QY 61 DRI-----FUNKPE 69
DB 145 DRIAPPLEVFVCPSSNSFDOVPE 169

```

```

RESULT 13
Q9ZOC4 O9ZOC4 PRELIMINARY; PRT; 208 AA.
AC Q9ZOC4;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DE 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE C1 AND C4 GENES, CLONE ABURA3-1, PARTIAL.
DE AND COMPLETE CDS (FRAGMENT).
GN Clacaco leaf curl virus.
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.
DR PRINTS: PR00228; GEMCOATCLVL1.
FT NON_TER 204 208
FT NON_TER 204 208
SQ SEQUENCE 208 AA; 23526 MW; 249CC31D8729C72D CRC64;

```

```

Query Match 69.2%; Score 247; DB 12; Length 208;
Best Local Similarity 58.8%; Pred. No. 3.9e-20;
Matches 50; Conservative 8; Mismatches 11; Indels 16; Gaps 1;

QY 1 TLVGCFFVQDGRSNGCGTSDNAAAEALNASSKEEALQIIRKIPKYLFAFALNSNL 60
DB 102 TLEMGTFVQDGRSNGCGTSDNAAAEALNASSKEEALQIIRKIPKYLFAFALNSNL 161
QY 61 DRI-----FUNKPE 69
DB 162 DRIAPPLEVFVCPSSNSFDOVPE 186

```

```

Query Match 69.2%; Score 247; DB 12; Length 208;
Best Local Similarity 58.8%; Pred. No. 3.9e-20;
Matches 50; Conservative 8; Mismatches 11; Indels 16; Gaps 1;

QY 1 TLVGCFFVQDGRSNGCGTSDNAAAEALNASSKEEALQIIRKIPKYLFAFALNSNL 60
DB 102 TLEMGTFVQDGRSNGCGTSDNAAAEALNASSKEEALQIIRKIPKYLFAFALNSNL 161
QY 61 DRI-----FUNKPE 69
DB 162 DRIAPPLEVFVCPSSNSFDOVPE 186

```

```

RESULT 14
Q9ZOC0 O9ZOC0 PRELIMINARY; PRT; 208 AA.
AC Q9ZOC0;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DE 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE C1 AND C4 GENES, CLONE ANG-1(B152),
DE PARTIAL AND COMPLETE CDS (FRAGMENT).
GN Clacaco leaf curl virus
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.
DR PRINTS: PR00228; GEMCOATCLVL1.
FT NON_TER 208 208
FT NON_TER 208 208
SQ SEQUENCE 208 AA; 23486 MW; E301135F799C3DAD CRC64;

```

```

Query Match 68.6%; Score 245; DB 12; Length 208;
Best Local Similarity 57.6%; Pred. No. 6.5e-20;
Matches 49; Conservative 9; Mismatches 11; Indels 16; Gaps 1;

QY 1 TLVGCFFVQDGRSNGCGTSDNAAAEALNASSKEEALQIIRKIPKYLFAFALNSNL 60
DB 102 TLEMGTFVQDGRSNGCGTSDNAAAEALNASSKEEALQIIRKIPKYLFAFALNSNL 161
QY 61 DRI-----FUNKPE 69
DB 162 DRIAPPLEVFVCPSSNSFDOVPE 186

```

```

RESULT 15
Q9ZOB8 Q9ZOB8 PRELIMINARY; PRT; 208 AA.
AC Q9ZOB8;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DE 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE C1 AND C4 GENES, CLONE ANG-2(B154),
DE PARTIAL AND COMPLETE CDS (FRAGMENT).
GN Clacaco leaf curl virus.
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.
DR PRINTS: PR00228; GEMCOATCLVL1.
FT NON_TER 204 208
FT NON_TER 204 208
SQ SEQUENCE 208 AA; 23472 MW; 629D0D0EF7C956AA CRC64;

```

```

Query Match 69.2%; Score 247; DB 12; Length 208;
Best Local Similarity 58.8%; Pred. No. 3.9e-20;
Matches 50; Conservative 8; Mismatches 11; Indels 16; Gaps 1;

QY 1 TLVGCFFVQDGRSNGCGTSDNAAAEALNASSKEEALQIIRKIPKYLFAFALNSNL 60
DB 102 TLEMGTFVQDGRSNGCGTSDNAAAEALNASSKEEALQIIRKIPKYLFAFALNSNL 161
QY 61 DRI-----FUNKPE 69
DB 162 DRIAPPLEVFVCPSSNSFDOVPE 186

```

Query Watch 68 68: Score 245; DB 12; Length 208:  
Pair Local Similarity 57.6%; Pct Ident 21.6%;  
Matches 49; Conservative 9; Mismatches 11; Indels 16; Gaps 1;  
QY 1 TLVNGEFOVDRSARGGCGQTSNFAAAALNASSKBEALQIREKIPKYLFAFAALNSNL 60  
DB 102 TLENGTFQIDGRSARGGCGQNDACAALNASSKADALATIREKLKDFIQVHNLNSNL 161  
QY 61 DRI-----FWKYPE 69  
DB 162 DRIFAPPLEVFCPPSSSFDQVPE 186

Search completed: February 3, 2001, 02:22:51  
Job time: 1888 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: February 3, 2001, 02:24:27 : Search time 83.07 seconds  
(without alignments)  
27.213 Million cell updates/sec

Title: us-09-289-346a-6

Percent score: 95.7

Sequence: 1 TLVWGFQVDCSRGSGCQT.....FAFAALNSLDRFDKTEP 70

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	340	95.2	352	1 VAL1_TGMV	P03567 tomato gold
2	240	67.2	361	1 VAL1_PYMV	P27258 potato yell
3	227	63.6	358	1 VAL1_CDKR	P14982 cassava lat
4	220	61.6	352	1 VAL1_TYLA	P14972 cassava lat
5	220	61.6	352	1 VAL1_TYLA	P14972 cassava lat
6	217	60.8	359	1 VAL1_TYLC	P38509 tomato yell
7	210	58.8	359	1 VAL1_TYLC	P38509 tomato yell
8	209	58.5	349	1 VAL1_PHVU	P06923 pepper huas
9	204	57.1	353	1 VAL1_BGMV	P05175 bean golden
10	203	56.9	355	1 VAL1_ABMY	P21947 abutilon mo
11	200	56.0	358	1 VAL1_BGTV	P14991 beet curly
12	199	55.0	358	1 VAL1_BGTV	P14991 beet curly
13	191	53.0	357	1 VAL1_TYLV	P27059 tomato yell
14	104	29.1	347	1 VAL1_SLVC	P25048 squash leaf
15	64.5	18.1	1713	1 LMA3_HUMAN	Q16787 homo sapien
16	63.5	17.8	1610	1 CCAD_MESAU	Q09244 mesocricetu
17	63.5	17.8	1610	1 CCAD_HUMAN	Q09244 mesocricetu
18	63.5	17.8	2203	1 CCAD_RAT	P27732 rattus norv
19	60	16.8	259	1 ONN_YEAST	P34964 saccharomy
20	60	16.8	259	1 ONN_YEAST	P34964 saccharomy
21	60	16.8	355	1 CPTB_RHOS	P14905 tobacco yell
22	59.5	16.7	299	1 Y175_HLAI	Q92407 helicobacte
23	59.5	16.7	470	1 RHSA_RHME	Q92407 helicobacte
24	59.5	16.7	4385	1 YP73_CAEEL	Q09222 caenorhabd
25	58.5	16.4	706	1 Y006_FICPL	Q92466 rickettsia
26	58.5	16.4	1232	1 Y005_CAEEL	Q92466 rickettsia
27	58	16.2	207	1 H16_MARMO	O85736 marimota mon
28	58	16.2	207	1 H16_MARMO	O85736 marimota mon
29	57.5	16.1	367	1 LHX4_MOUSE	P53776 phosphatidyl
30	57.5	16.1	844	1 SECA_STYCA	P47994 straphylococ
31	57.5	16.1	1852	1 CCAS_CYPCA	P22316 cyprinus ca
32	57	16.0	1483	1 CYP1_YEAST	P12351 saccharomy
33	57	16.0	1608	1 HLYA_SMRMA	P15320 saccharia ma

34 56.5 15-8 2150 1 CCAD\_CHICK  
35 56 15-7 258 1 YL48\_MYCTU  
36 56 15-7 408 1 CINA\_THIEM  
37 56 15-7 703 1 UGS21\_RAT  
38 56 15-7 1034 1 Y001\_MOUSE  
39 56 15-7 1034 1 Y001\_MOUSE  
40 56 15-7 1031 1 CARB\_SULSO  
41 56 15-7 1663 1 HAPD\_HUMAN  
42 55.5 15-5 266 1 ETFFB\_MYCLE  
43 55.5 15-5 266 1 ETFFB\_MYCTU  
44 55.5 15-5 419 1 Y223\_MYCPN  
45 55.5 15-5 967 1 HEPA\_ECOLI

## ALIGNMENTS

RESULT 1  
VAL1\_TGMV STANDARD; PRT; 352 AA.  
ID VAL1\_TGMV  
AC C  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DE 01-JUN-1994 (Rel. 29, Last annotation update)  
DE AL1 PROTEIN.  
GN ACL.

OS Tomato golden mosaic virus (TGMV).  
OC Viruses: ssDNA viruses: Geminiviridae; Begomovirus.  
RN  
PR  
SEQUENCE FROM N.A.

RA Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.  
RT "Complete nucleotide sequence of the infectious cloned DNA components  
of tomato golden mosaic virus: potential coding regions and regulatory  
sequences.";

RL EMO J. 3:2197-2205(1984).

CC -! SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/annouce/  
or send an email to license@sib-sib.ch).

CC EMBL: K02029; -! NOT\_ANNOTATED\_CDS.  
DR PIR: A04170; Q0CVL1  
DR INTERPRO: IPR001191  
DR INTERPRO: IPR001301  
DR PFAM: PF00759; Gemin1-AL1; 1.  
DR PRINTS: PR00227; GEMCONTAL1.  
DR PROSITE: PS00228; GEMCONTAL1.  
DR ATTC: ATTC0228; GEMCONTAL1.  
FT NP\_BIND 223 230 ATP (POTENTIAL).

SQ SEQUENCE 352 AA; 40332 MW; C33C9386964B4A CRC64;

Query Match 95.2%; Score 340; DB 1; Length 352;

Best Local Similarity 95.7%; Pred. No. 7.8e-31;

Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 TLVWGFQVDCSRGSGCQTNDAAAFALNSKEEALQIIRKTPKYLFAFAALNSL 60

Db 111 TLVWGFQVDCSRGSGCQTNDAAAFALNSKEEALQIIRKTPKYLFAFAALNSL 170

Oy 61 DRFDKTEP 70

Db 171 DRFDKTEP 180

RESULT 2

VAL1\_PYMV STANDARD; PRT; 361 AA.

```

P27258;
AT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE ALU PROTEIN (Rel. 23, Last annotation update)
OS Potato yellow mosaic virus (isolate Venezuela).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
R1 SEQUENCE FROM N.A.
RX MEDLINE-91311403; PubMed-1856690;
RA Cuddeles R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.:
RT "Nucleotide sequence of the infectious cloned DNA components of
RT potato yellow mosaic virus (strain West Konyan 844)."
RL J. Gen. Virol. 72:1515-1520(1991).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
DR EMBL; D00940; BAA00782.1; -
DR PIR; J00364; QOCVPT.
DR INTERPRO; IPR001301; -
DR PPAM; PF00759; Gemin_A1; 1.
DR PRINTS; PR00227; GEMCOATL1.
DR PRINTS; PR00228; GEMCOATCLV1.
KW ATP-binding.
ATP BINDING. 229
FT NP_BIND 361 AA: 40950 MW; 5627A33BF1264383 CRC64;
SQ SEQUENCE 361 AA: 40950 MW; 5627A33BF1264383 CRC64;

Query Match 57.2% Score 240; DB 1; Length 361;
Best Local Similarity 66.7% Pred. No. 1,2e-19;
Matches 46; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDSRGSGGQTSNDAAAFALNASKEEALQIREKIPERYLFAALNSNL 60
DB 110 TLVWGEFQVDSRGSGGQTSNDAAAFALNASKEEALQIREKIPERYLFAALNSNL 169
QY 61 DRIFDKPTEP 69
DB 170 DRIFDKAPE 178

Query Match 57.2% Score 240; DB 1; Length 361;
Best Local Similarity 66.7% Pred. No. 1,2e-19;
Matches 46; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDSRGSGGQTSNDAAAFALNASKEEALQIREKIPERYLFAALNSNL 60
DB 110 TLVWGEFQVDSRGSGGQTSNDAAAFALNASKEEALQIREKIPERYLFAALNSNL 169
QY 61 DRIFDKPTEP 69
DB 170 DRIFDKAPE 178

RESULT 3
ID VALL CLVX STANDARD; PRT; 358 AA.
AC P14982;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DE ALU PROTEIN (40.4 KDA Protein)
GN AC1. Cassava latent virus (strain West Konyan 844).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
R1 SEQUENCE FROM N.A.
RX MEDLINE-90174930; PubMed-2308831;
RA Morris B., Coates L., Lowe S., Richardson K., Eddy P.:
RT "Nucleotide sequence of the infectious cloned DNA components of
RT African cassava mosaic virus (strain Nigerian 1990)."
RL J. Gen. Virol. 71:1979-1986(1990).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
DR EMBL; X17055; CAA34953.1; -
DR PIR; S07594; S07594.
DR INTERPRO; IPR001191; -
DR PPAM; PF001301; -
DR PRINTS; PR00227; GEMCOATL1.
DR PRINTS; PR00228; GEMCOATCLV1.
KW ATP-binding.
ATP BINDING. 220
FT NP_BIND 358 AA: 40435 MW; 1DB16B80CB2D5E2C CRC64;
SQ SEQUENCE 358 AA: 40435 MW; 1DB16B80CB2D5E2C CRC64;

Query Match 63.6% Score 227; DB 1; Length 358;
Best Local Similarity 60.0% Pred. No. 3,4e-18;
Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDSRGSGGQTSNDAAAFALNASKEEALQIREKIPERYLFAALNSNL 60
DB 109 TLVWGEFQVDSRGSGGQTSNDAAAFALNASKEEALQIREKIPERYLFAALNSNL 168
QY 61 DRIFDKPTEP 70
DB 169 DRIFQEPAP 178

```

```

CC or send an email to license@isb-sib.ch)
CC
DR EMBL; J02057; -; NOT_ANNOTATED_CDS.
DR INTERPRO; IPR001301; -
DR PPAM; PF00759; Gemin_A1; 1.
DR PRINTS; PR00227; GEMCOATL1.
DR PRINTS; PR00228; GEMCOATCLV1.
KW ATP-binding.
ATP BINDING. 220
FT NP_BIND 358 AA: 40346 MW; ED173E753EE92D49 CRC64;
SQ SEQUENCE 358 AA: 40346 MW; ED173E753EE92D49 CRC64;

Query Match 63.6% Score 227; DB 1; Length 358;
Best Local Similarity 60.0% Pred. No. 3,4e-18;
Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDSRGSGGQTSNDAAAFALNASKEEALQIREKIPERYLFAALNSNL 60
DB 109 TLVWGEFQVDSRGSGGQTSNDAAAFALNASKEEALQIREKIPERYLFAALNSNL 168
QY 61 DRIFDKPTEP 70
DB 169 DRIFQEPAP 178

RESULT 4
ID VALL CLVX STANDARD; PRT; 358 AA.
AC P14972;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DE ALU PROTEIN (40.4 KDA Protein)
GN AC1. Cassava latent virus (strain Nigerian)
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
R1 SEQUENCE FROM N.A.
RX MEDLINE-90174930; PubMed-2308831;
RA Morris B., Coates L., Lowe S., Richardson K., Eddy P.:
RT "Nucleotide sequence of the infectious cloned DNA components of
RT African cassava mosaic virus (strain Nigerian 1990)."
RL J. Gen. Virol. 71:1979-1986(1990).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
DR EMBL; X17055; CAA34953.1; -
DR PIR; S07594; S07594.
DR INTERPRO; IPR001191; -
DR PPAM; PF001301; -
DR PRINTS; PR00227; GEMCOATL1.
DR PRINTS; PR00228; GEMCOATCLV1.
KW ATP-binding.
ATP BINDING. 220
FT NP_BIND 358 AA: 40435 MW; 1DB16B80CB2D5E2C CRC64;
SQ SEQUENCE 358 AA: 40435 MW; 1DB16B80CB2D5E2C CRC64;

Query Match 63.6% Score 227; DB 1; Length 358;
Best Local Similarity 60.0% Pred. No. 3,4e-18;
Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDSRGSGGQTSNDAAAFALNASKEEALQIREKIPERYLFAALNSNL 60
DB 109 TLVWGEFQVDSRGSGGQTSNDAAAFALNASKEEALQIREKIPERYLFAALNSNL 168

```



```

Oy 61 DRIFDXTPEP 70
Db 169 DRIFQFPAP 178

RESULT 5
ID VAL1_TYLCU STANDARD: PRT: 362 AA.
AC P36279;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN (C1 PROTEIN).
GN C1.
OS Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP MEDLINE=94256836; PubMed=8198442;
EX DRY I.B., Sigden J.R., Krake L.R., Mullineaux P.M., Rezaian M.A.;
RA "Nucleotide sequence and genome organization of tomato leaf curl
RT geminivirus.";
RL J. Gen. Virol. 74:147-151(1993).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: S52593; GenBank: AF0001191.
DR INTERPRO: IPR001191;
DR PIR: S39211; S39211.
DR PFAM: PF00759; Geminol_1;
DR PRINTS: PR00227; GEMCOATV1.
DR PROSITE: PR00228; GEMCOATCLVL1.
KW ATP-binding. 221 228 ATP (POTENTIAL).
FT NP_BIND 359 AA: 41065 MW; 20170A51EF80A3EC CRC64;
SQ SEQUENCE 362 AA: 41197 MW; 343E7184E4704098 CRC64;

Query Match 61.6%; Score 220; DB 1; Length 362;
Best Local Similarity 52.9%; Pred. No. 2.1e-17;
Matches 43; Conservative 9; Mismatches 15; Indels 16; Gaps 1;

Oy 1 TLVGFQVDGRSGRGCGQTNDAAEALNASSKEALQIREKIPKYLFAFALNSML 60
Db 110 TLENGEFQDGRSGRGCGQTNDAAEALNASSKEALQIREKIPKYLFAFALNSML 169

Oy 61 DRI
Db 170 DRIFTPPEVNSPFLSSSFQVPE 194

RESULT 6
ID VAL1_TYLCU STANDARD: PRT: 359 AA.
AC P38609;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE AL1 PROTEIN (C1 PROTEIN).
GN C1.
OS Tomato yellow leaf curl virus (strain Murcia) (TYLCV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP MEDLINE=94256836; PubMed=8198442;

Noris E., Hidalgo E., Accotto G., Moriones E.;
RT "High similarity among the tomato yellow leaf curl virus isolates
RT from the west Mediterranean basin: the nucleotide sequence of an
RT infectious clone from Spain;
RT Arch. Virol. 135:193-194;
CC SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: Z25751; GenBank: AF0001191.
DR INTERPRO: IPR001191;
DR PIR: S39211; S39211.
DR PFAM: PF00759; Geminol_1;
DR PRINTS: PR00227; GEMCOATV1.
DR PROSITE: PR00228; GEMCOATCLVL1.
KW ATP-binding. 221 228 ATP (POTENTIAL).
FT NP_BIND 359 AA: 41065 MW; 20170A51EF80A3EC CRC64;
SQ SEQUENCE 359 AA: 41065 MW; 20170A51EF80A3EC CRC64;

Query Match 50.8%; Score 217; DB 1; Length 359;
Best Local Similarity 52.4%; Pred. No. 4.5e-17;
Matches 41; Conservative 9; Mismatches 19; Indels 0; Gaps 0;

Oy 2 LVWGFQVDGRSGRGCGQTNDAAEALNASSKEALQIREKIPKYLFAFALNSML 61
Db 111 LEWGFQDGRSGRGCGQTNDAAEALNASSKEALQIREKIPKYLFAFALNSML 170

Oy 62 RIFDKTPEP 70
Db 171 RVFQVPPAP 179

RESULT 7
ID VAL1_TYLCU STANDARD: PRT: 359 AA.
AC P38609;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN (C1 PROTEIN).
GN C1.
OS Tomato yellow leaf curl virus (strain Marmande) (TYLCV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP MEDLINE=92107650; PubMed=1840676;
EX Kheyri-Pour A., Bendahmane M., Matzeit V., Accotto G.P., Crespi S.,
RA Gronenborn B.;
RT "Tomato yellow leaf curl virus from Sardinia is a
RT whitefly-transmitted monopartite geminivirus.";
RL Nucleic Acids Res. 19:5763-5769(1991).
CC SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: X61153; GenBank: AF0001191.
DR INTERPRO: IPR001191;
DR PIR: S22593; S22593.
DR PFAM: PF00759; Geminol_1;
DR PRINTS: PR00227; GEMCOATV1.
DR PROSITE: PR00228; GEMCOATCLVL1.
KW ATP-binding. 221 228 ATP (POTENTIAL).
FT NP_BIND 359 AA: 41065 MW; 20170A51EF80A3EC CRC64;
SQ SEQUENCE 359 AA: 41065 MW; 20170A51EF80A3EC CRC64;

```

DR PFAM: PF00759; Gemin1-ALL; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR PRINTS: PR00228; GEMCOATCLVL.  
 DR KW ATP-binding.  
 DR FT NP-BIND 220 227  
 DR SQ SEQUENCE 359 AA; 40733 MW; 9717B407C33EAF7 CRC64;  
 Query Match 58.8%; Score 210; DB 1; Length 359;  
 Best Local Similarity 56.5%; Pred. No. 2.8e-16;  
 Matches 39; Conservative 11; Mismatches 10; Gaps 0;  
 QY 2 LWGEFQVDSRGSGCOTSDNDAAEALNASSKEEALQIREKIPKYLPAFAPALNSLD 61  
 DB 111 LWMGFQIGDSRGSGCOTSDNDAAEALNASSKEEALQIREKIPKYLPAFAPALNSLD 170  
 QY 62 RFDKTPPEP 70  
 DB 171 KVFQVPPAP 179  
 RESULT 8  
 ID VAL1-ABWV STANDARD; PRT; 349 AA.  
 AC 006923;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE ALL PROTEIN.  
 GN AL1.  
 CC This Swiss-Prot entry copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: X70418; CAA49856.1; .  
 CC PIR: S31875; S31875.  
 CC PIR: QJ2300; QJ2300.  
 CC INTERPRO: IPR001191; .  
 CC PFAM: PF00759; Gemin1-ALL; 1.  
 CC PRINTS: PR00227; GEMCOATALL.  
 CC PRINTS: PR00228; GEMCOATCLVL.  
 CC ATP-binding.  
 CC NP-BIND 221 228  
 CC SQ SEQUENCE 349 AA; 39722 MW; D3F4E76CD56370F4 CRC64;  
 Query Match 58.5%; Score 209; DB 1; Length 349;  
 Best Local Similarity 57.1%; Pred. No. 3.5e-16;  
 Matches 40; Conservative 13; Mismatches 17; Indels 0; Gaps 0;  
 QY 1 TLWGEFQVDSRGSGCOTSDNDAAEALNASSKEEALQIREKIPKYLPAFAPALNSL 60  
 DB 110 TVNGEFQVDSRGSGCOTSDNDAAEALNASSKEEALQIREKIPKYLPAFAPALNSL 169  
 QY 61 DRFDKTPPEP 70  
 DB 170 ERIFKVPPEP 179  
 RESULT 10  
 ID VAL1-ABWV STANDARD; PRT; 355 AA.  
 AC 021947;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE ALL PROTEIN.  
 GN AL1.  
 CC This Swiss-Prot entry copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: X70418; CAA49856.1; .  
 CC PIR: S31875; S31875.  
 CC PIR: QJ2300; QJ2300.  
 CC INTERPRO: IPR001191; .  
 CC PFAM: PF00759; Gemin1-ALL; 1.  
 CC PRINTS: PR00227; GEMCOATALL.  
 CC PRINTS: PR00228; GEMCOATCLVL.  
 CC ATP-binding.  
 CC NP-BIND 221 228  
 CC SQ SEQUENCE 349 AA; 39722 MW; D3F4E76CD56370F4 CRC64;  
 Query Match 57.1%; Score 204; DB 1; Length 353;  
 Best Local Similarity 57.1%; Pred. No. 1.3e-15;  
 Matches 40; Conservative 13; Mismatches 17; Indels 0; Gaps 0;  
 QY 1 TLWGEFQVDSRGSGCOTSDNDAAEALNASSKEEALQIREKIPKYLPAFAPALNSL 60  
 DB 110 TVNGEFQVDSRGSGCOTSDNDAAEALNASSKEEALQIREKIPKYLPAFAPALNSL 169  
 QY 61 DRFDKTPPEP 70  
 DB 170 ERIFKVPPEP 179  
 RESULT 9  
 ID VAL1-ABWV STANDARD; PRT; 353 AA.  
 AC 006923;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE ALL PROTEIN (4.2 KDA PROTEIN).  
 GN AC1.  
 CC Bean golden mosaic virus.  
 CC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 CC SEQUENCE FROM N.A.  
 CC Howarth A.J., Caton J., Bossert M., Goodman R.M.;  
 CC "Nucleotide sequence of bean golden mosaic virus and a model for gene  
 CC regulation in geminiviruses.";  
 CC Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).  
 CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC This Swiss-Prot entry copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: M10070; M10070.  
 CC PIR: S31875; S31875.  
 CC INTERPRO: IPR001191; .  
 CC PFAM: PF00759; Gemin1-ALL; 1.  
 CC PRINTS: PR00227; GEMCOATALL.  
 CC PRINTS: PR00228; GEMCOATCLVL.  
 CC ATP-binding.  
 CC NP-BIND 222 229  
 CC SQ SEQUENCE 353 AA; 40190 MW; 80FA750F6029A34 CRC64;  
 Query Match 57.1%; Score 204; DB 1; Length 353;  
 Best Local Similarity 57.1%; Pred. No. 1.3e-15;  
 Matches 40; Conservative 13; Mismatches 17; Indels 0; Gaps 0;  
 QY 1 TLWGEFQVDSRGSGCOTSDNDAAEALNASSKEEALQIREKIPKYLPAFAPALNSL 60  
 DB 110 TVNGEFQVDSRGSGCOTSDNDAAEALNASSKEEALQIREKIPKYLPAFAPALNSL 169  
 QY 61 DRFDKTPPEP 70  
 DB 170 ERIFKVPPEP 179  
 RESULT 10  
 ID VAL1-ABWV STANDARD; PRT; 355 AA.  
 AC 021947;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE ALL PROTEIN.  
 GN AL1.  
 CC This Swiss-Prot entry copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: X70418; CAA49856.1; .  
 CC PIR: S31875; S31875.  
 CC PIR: QJ2300; QJ2300.  
 CC INTERPRO: IPR001191; .  
 CC PFAM: PF00759; Gemin1-ALL; 1.  
 CC PRINTS: PR00227; GEMCOATALL.  
 CC PRINTS: PR00228; GEMCOATCLVL.  
 CC ATP-binding.  
 CC NP-BIND 221 228  
 CC SQ SEQUENCE 349 AA; 39722 MW; D3F4E76CD56370F4 CRC64;  
 Query Match 58.5%; Score 209; DB 1; Length 349;  
 Best Local Similarity 57.1%; Pred. No. 3.5e-16;  
 Matches 40; Conservative 13; Mismatches 17; Indels 0; Gaps 0;  
 QY 1 TLWGEFQVDSRGSGCOTSDNDAAEALNASSKEEALQIREKIPKYLPAFAPALNSL 60  
 DB 110 TVNGEFQVDSRGSGCOTSDNDAAEALNASSKEEALQIREKIPKYLPAFAPALNSL 169  
 QY 61 DRFDKTPPEP 70  
 DB 170 ERIFKVPPEP 179  
 RESULT 10  
 ID VAL1-ABWV STANDARD; PRT; 355 AA.  
 AC 021947;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE ALL PROTEIN.  
 GN AL1.  
 CC This Swiss-Prot entry copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: X70418; CAA49856.1; .  
 CC PIR: S31875; S31875.  
 CC PIR: QJ2300; QJ2300.  
 CC INTERPRO: IPR001191; .  
 CC PFAM: PF00759; Gemin1-ALL; 1.  
 CC PRINTS: PR00227; GEMCOATALL.  
 CC PRINTS: PR00228; GEMCOATCLVL.  
 CC ATP-binding.  
 CC NP-BIND 221 228  
 CC SQ SEQUENCE 349 AA; 39722 MW; D3F4E76CD56370F4 CRC64;  
 Query Match 58.5%; Score 209; DB 1; Length 349;  
 Best Local Similarity 57.1%; Pred. No. 3.5e-16;  
 Matches 40; Conservative 13; Mismatches 17; Indels 0; Gaps 0;



```

OS Tomato yellow leaf curl virus (TYLCV).
RN Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RX MEDLINE-92024070. PubMed-1926771.
RA Navot N., Pichersky E., Zeldan M., Zamir D., Czosnek H.:
RT "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
RL Virology 185:151-151(1991).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outposts
CC at the European Bioinformatics Institute. The data are made available
CC for use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC QY 4 WGEFVQDSRGAGCCGTSNDAAEALNASKKEALQIREKIPKYLEFAALNSNDRIF 63
CC DB 111 FGVSDIDCRSARGCCGTSNDAAEALNASKKEALQIREKIPKYLEFAALNSNDRIF 170
CC QY 64 F 64
CC DB 171 F 171
CC
CC RESULT 14
CC VAL1 SLICV STANDARD: PRT: 347 AA.
CC AC P29048:
CC DT 01-DEC-1992 (Rel. 24, Created)
CC DT 01-DEC-1992 (Rel. 24, Last sequence update)
CC DE ALL PROTEIN (Rel. 24, Last annotation update)
CC OS Squash leaf curl virus.
CC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
CC RN [1]
CC SEQUENCE FROM N.A.
CC RX MEDLINE-91082449; PubMed-1984668;
CC RA Mazarowitz S.G., Lezidin I.B.:
CC RT "Components of a bipartite squash leaf curl geminivirus with a broad
CC host range phenotype."
CC RL Virology 180:58-69(1991).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outposts
CC at the European Bioinformatics Institute. The data are made available
CC for use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DB EMBL; M38183; AC034410.1; ALT-INIT.

```

```

DR PIR: C36785; QOCV51.
DR INTERPRO: IPR001191: -.
DR FRAM: PF00799; Gemini_All: 1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00228; GEMCOATL1.1.
KW ATP-binding, 218 225 ATP (POTENTIAL).
FT NP_BIND 347 AA: AFDAGBDD122110E CRC64;
SQ SEQUENCE 347 AA: AFDAGBDD122110E CRC64;
-----
Query Match 29.18; Score 104; DB 1; Length 347;
Best Local Similarity 34.88; Pred. No. 0.00015; Indels 4; Gaps 1;
Matches 23; Conservative 14; Mismatches 25;
QY 5 GEFVQDSRGAGCCGTSNDAAEALNASKKEALQIREKIPKYLEFAALNSNDRIF 64
DB 116 GQTKVSG---GSKSNKQDVYHNAVNGSAGEALDIKAGDKPTFIVYHNLNVERLF 171
QY 65 DKTPPE 70
DB 172 QKTPPE 177
-----
RESULT 15
LMA3 HUMAN STANDARD: PRT: 1713 AA.
ID LMA3_HUMAN
DC Q16787; Q16787; Q13680 (Created)
RX Ryan M.C., Fizard R., Vandevanter D.R., Carter W.G.:
RT "Cloning of the Lama3 gene encoding the alpha 3 chain of the adhesive
RT ligand epiligrin. Expression in wound repair."
RN J Biol. Chem. 269:22779-22787(1994).
RP SEQUENCE OF 1-1331 FROM N.A. (ISOPFORMS A AND B).
RX MEDLINE-94357926; PubMed-8077230:
RA Ryan M.C., Fizard R., Vandevanter D.R., Carter W.G.:
RT "Cloning of the Lama3 gene encoding the alpha 3 chain of the adhesive
RT ligand epiligrin. Expression in wound repair."
RN J Biol. Chem. 269:22779-22787(1994).
RP SEQUENCE OF 1-1331 FROM N.A. (ISOPFORMS A AND B).
RX MEDLINE-96163880; PubMed-8586427:
RA Vidal F., Baudouin C., Miquel C., Galliano M.-F., Christiano A.M.,
RA Utito J., Ortonne J.-P., Meneguizzi G.:
RT "Cloning of the laminin alpha 3 chain gene (LAMA3) and identification
RT of a homozygous deletion in a patient with Herlitz junctional
RL Epidermolysis bullosa."
CC -1- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
CC IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT.
CC -1- WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
CC -1- FUNCTION: LAMININ-5 IS THOUGHT TO BE INVOLVED IN (1) CELL ADHESION
CC & CELL MIGRATION, (2) CELL ADHESION AND CELL ADHESION TO EXTRACELLULAR
CC MATRIX IN HEMIDESMOSES, (3) CELL ADHESION AND CELL ADHESION TO
CC PHOSPHORYLATION OF PP125-PAK AND P80, (3) DIFFERENTIATION OF
CC KERATINOCYTES.
CC -1- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
CC THE LONG ARM OF THE MOLECULE IS LAMININ-5 (EPILIGRIN/LAMININ/
CC NICLIN) AND POSSIBLY ALSO A COMPONENT OF LAMININ-6 (N-LAMININ)
CC AND LAMININ-7 (KS-LAMININ).
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT
CC MEMBRANES (MAJOR COMPONENT).
CC -1- ALTERNATIVE PRODUCTS: THE TWO ISOFORMS A AND B DIFFER IN THEIR N-
CC TERMINUS. THE SEQUENCE SHOWN HERE IS THAT OF THE SMALLER VARIANT

```

CC CC -!- TISSUE SPECIFICITY: SKIN; RESPIRATORY, URINARY, AND DIGESTIVE  
 CC CC -!- CELL AND CELLULAR FUNCTIONS WITH PROMINENT  
 CC CC SECRETORY OR PROTECTIVE FUNCTIONS IN EPIDERMIS,  
 CC CC AND EPITHELIAL CELL TONGUE THAT MIGRATES INTO A WOUND BED, A  
 CC CC DIFFERENTIAL AND FOCAL EXPRESSION OF THE ALPHA-3 CHAIN IS OBSERVED  
 CC CC IN THE CNS.  
 CC CC -!- INDUCTION: LAMININ-5 IS UP-REGULATED IN WOUND SITES OF HUMAN SKIN.  
 CC CC -!- DOMAIN: THE ALPHA-HELICAL DOMAINS 1 AND IT ARE THOUGHT TO INTERACT  
 CC CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.  
 CC CC -!- DISEASE: JUNCTIONAL EPIDERMOLYSIS BULLOSA GRAVIS IS A BLISTERING  
 CC CC DISORDER IN SKIN THAT IS CHARACTERIZED BY A SEPARATION OF BASAL  
 CC CC CELLS FROM THE BASEMENT MEMBRANE DUE TO A DECREASED NUMBER OF  
 CC CC HEMIDESMOSOMES. LAMININ-5 IS MISSING FROM THE BASEMENT MEMBRANE OF  
 CC CC PATIENTS WITH THE GRAVIS FORM OF EPIDERMOLYSIS BULLOSA.  
 CC CC -!- SIMILARITY: CONTAINS 2.5 LAMININ EGF-LIKE DOMAINS.  
 CC CC -!- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.  
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation.  
 CC CC The European Bioinformatics Institute. There are no restrictions on its  
 CC CC use by non-profit institutions as long as its content is in no way  
 CC CC modified and this statement is not removed. Usage by and for commercial  
 CC CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC CC EMBL: L34155; AA59483.1; ;  
 CC CC EMBL: X85107; CA59428.1; ;  
 CC CC EMBL: X85108; CA59429.1; ;  
 CC CC DR HSSP; P02468; 1TLE.  
 CC CC DR MIR: 600805; ;  
 CC CC DR MIR: 226700; ;  
 CC CC DR INTERPRO; IPRO00561; ;  
 CC CC DR INTERPRO; IPRO02049; ;  
 CC CC DR PFAM: PF00053; laminin\_EGF; 2.  
 CC CC DR PFAM: PF00054; laminin\_G; 2.  
 CC CC DR PROSITE: PS00022; EGF\_1; 1.  
 CC CC DR PROSITE: PS01186; EGF\_2; 1.  
 CC CC DR GlycoProtein; Basement membrane; Extracellular matrix; Coiled coil;  
 CC CC Glycoprotein; Cell adhesion; Cell growth; Cell death; Cell adhesion; Signal;  
 CC KW Alternative splicing; Cell adhesion; repeat; Signal;  
 CC SIGNAL: 1 20  
 CC FT CHAIN POTENTIAL.  
 CC FT LAMININ ALPHA-3 CHAIN.  
 CC FT DOMAIN 21 1713  
 CC FT DOMAIN 46 201  
 CC FT DOMAIN 67 185  
 CC FT DOMAIN 67 185  
 CC FT DOMAIN 67 185  
 CC FT DOMAIN 114 166  
 CC FT DOMAIN 114 166  
 CC FT DOMAIN 202 793  
 CC FT DOMAIN 202 793  
 CC FT DOMAIN 794 1713  
 CC FT DOMAIN 794 1713  
 CC FT DOMAIN 794 1713  
 CC FT DOMAIN 971 1139  
 CC FT DOMAIN 1140 1353  
 CC FT DOMAIN 1354 1529  
 CC FT DOMAIN 1530 1713  
 CC FT DOMAIN 356 548  
 CC FT DOMAIN 356 548  
 CC FT DOMAIN 594 621  
 CC FT DOMAIN 702 765  
 CC FT DOMAIN 1686 1713  
 CC FT DOMAIN 67 76  
 CC FT DISULFID 67 76  
 CC FT DISULFID 69 83  
 CC FT DISULFID 98 115  
 CC FT DISULFID 98 115  
 CC FT DISULFID 114 126  
 CC FT DISULFID 116 135  
 CC FT DISULFID 137 146  
 CC FT DISULFID 149 164  
 CC FT DISULFID 202 202  
 CC FT DISULFID 205 205  
 CC FT DISULFID 205 205

FT SITE 658 660  
 FT CARBOHYD 542 542  
 FT CARBOHYD 545 545  
 FT CARBOHYD 545 545  
 FT CARBOHYD 882 882  
 FT CARBOHYD 882 882  
 FT CARBOHYD 964 964  
 FT CARBOHYD 1108 1108  
 FT CARBOHYD 1131 1131  
 FT CARBOHYD 1325 1325  
 FT CARBOHYD 1477 1477  
 FT CARBOHYD 1667 1667  
 FT VARSPLIC 1 45  
 FT CONFLICT 5 5  
 FT CONFLICT 123 125  
 FT CONFLICT 481 481  
 FT CONFLICT 481 481  
 FT CONFLICT 969 969  
 FT CONFLICT 1052 1052  
 FT CONFLICT 1052 1052  
 FT CONFLICT 1184 1184  
 FT SEQUENCE 1713 AA: 189304 MW: 45EA9BE1017B60D3 CRC64;  
 SQ  
 Query Match 18.1%; Score 64.5; DB 1; Length 1713;  
 Best Local Similarity 34.4%; Pred. No. 29;  
 Matches 21; Conservative 12; Mismatches 21; Indels 7; Gaps 3;  
 QY 9 VDHSGAGCGCTSDNDAAEALN--ASSKEEALQ-TIREKIPKYLEFAALNSNLDIFD 65  
 DB 553 VDAATAYENILNAIKAEADAANRAASASALOTVIKEDLPK---ARTLSNSDKLN 608  
 QY 66 K 66  
 DB 609 E 609

Search completed: February 3, 2001, 02:24:28  
 Job time: 640 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:17:39 ; Search time 118.74 Seconds  
(without alignments)  
40.029 Million cell updates/sec

Title: US-09-289-346a-6

Perfect score: 357

Sequence: 1 TLVNGFEVQDGRSGCGQT.....EAFEAALNSLDRIFDXTPEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR66:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result	No.	Score	Match	Length	DB	ID	Description
1	340	95.2	352	1	QCVCVL1		All protein - toma
2	240	67.2	1	QCVCPT			All protein - toma
3	227	63.6	358	2	S07594		hypothetical prote
4	220	61.6	362	1	J01887		All protein - toma
5	217	60.8	359	2	S59211		gene C1 protein -
6	212	59.4	321	2	J02327		All protein - indi
7	212	59.4	321	2	J02327		hypothetical prote
8	209	58.5	349	2	J02300		hypothetical prote
9	209	58.5	349	2	S31875		All protein - pep
10	204	57.1	359	2	S39235		gene C1 protein -
11	203	56.9	355	1	QCQCVW1		AV1 protein - abut
12	200	56.0	358	1	J01870		All protein - toma
13	200	56.0	385	2	S28360		All protein - beet
14	193	54.1	360	2	S59485		replication-associ
15	193	54.1	360	2	S59485		All protein - toma
16	118	33.1	1	QCQV59			All protein - clone
17	104	29.1	347	1	QCQV51		All protein - toma
18	67	16.8	447	2	T12544		hypothetical prote
19	64.5	18.1	587	2	JC1419		Fc gamma (Igc) rec
20	64.5	18.1	1713	2	A55347		adhesive ligand ep
21	63.5	17.8	1610	2	A46227		voltage-dependent
22	63.5	17.8	1646	2	JH0422		voltage-dependent
23	63.5	17.8	1646	2	JH0422		calcium channel al
24	63.5	17.8	2181	2	A58198		voltage-dependent
25	63.5	17.8	2203	2	T42742		voltage-dependent
26	63	17.6	316	2	C92085		conserved hypotet
27	61.5	17.2	392	2	T45290		acetyl-CoA C-acety
28	60.5	16.9	840	2	T36175		probable large ATP
29	60	16.8	269	2	S61633		YNT20 protein - ye

## ALIGNMENTS

RESULT 1

QCVCVL1

All protein - tomato golden mosaic virus

C:Species: tomato golden mosaic virus

A:Note: host Nicotiana sp. (tobacco)

C:Date: 28-Aug-1985 #sequence\_revision 28-Aug-1985 #text\_change 08-Apr-1994

C:Accession: A04170 Steijn, V.E.; Coutts, R.H.A.; Buck, K.W.

EMBL J\_3, 2197-2205, 1994

A:Title: Complete nucleotide sequence of the infectious cloned DNA components of toma

A:Reference number: A04163

A:Accession: A04170

A:Molecule type: DNA

A:Residues: 1-352 <HAM>

C:Comment: The genome consists of two circular, single-stranded DNA components, DNA A

C:Genetics:

A:Map position: segment A

C:Superfamily: tomato golden mosaic virus All protein

Query Match Best Local Similarity 95.2%; Score 340; DB 1; Length 352;

Hatches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TLVNGFEVQDGRSGCGQTNDAAALNASSKEALQIIREKPEKYLEFAALNSNL 60

Db 111 TLVNGFEVQDGRSGCGQTNDAAALNASSKEALQIIREKPEKYLEFAALNSNL 170

Qy 61 DRIFDXTPEP 70

Db 171 DRIFDXTPEP 180

RESULT 2

QCVCPT

All protein - potato yellow mosaic virus (isolate Venezuela)

C:Species: potato yellow mosaic virus

C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Jun-2000

C:Accession: J00362 J. Gen. Virol. 72, 1515-1520, 1991

B:Coutts, R.H.A.; Coffin, R.S.; Roberts, R.J.F.; Hamilton, W.D.O.

A:Title: The nucleotide sequence of the infectious cloned DNA components of potato ye

A:Reference number: J00362; MUID:91311403

A:Accession: J00364

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-352 <HAM>

C:Genetics:

A:Map position: segment A

C:Superfamily: tomato golden mosaic virus All protein

```

Query Match          67.2%; Score 240; DB 1; Length 361;
Best Local Similarity 66.7%; Pred. No. 9.6e-20;
Matches 46; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 1 TLWGEFQDGRSARGGCGTNDAAALNASSKEALQIIREKIPKYLFAFALNSNL 60
D 1 TLWGEFQDGRSARGGCGTNDAAALNASSKEALQIIREKIPKYLFAFALNSNL 60
D 11 TLWGEFQDGRSARGGCGTNDAAALNASSKEALQIIREKIPKYLFAFALNSNL 169

QY 61 DRIFDKTPEP 69
D 11 DRIFDKTPEP 178
D 170 DRIFDKTPEP 178

RESULT 3
S07594
hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)
C:Species: cassava latent virus
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
R:Morris, G.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.
Nucleic Acids Res. 18, 197-198, 1990
A:Title: Nucleotide sequence of the infectious cloned DNA components of African cassava
A:Reference number: S07590; MUID:90174930
A:Accession: S07594
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-362 <DR>
A:Cross-references: ENBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376
C:Genetics:
A:Map position: segment DNAL
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match          63.6%; Score 227; DB 2; Length 358;
Best Local Similarity 60.0%; Pred. No. 2.9e-18;
Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

QY 1 TLWGEFQDGRSARGGCGTNDAAALNASSKEALQIIREKIPKYLFAFALNSNL 60
D 1 TLWGEFQDGRSARGGCGTNDAAALNASSKEALQIIREKIPKYLFAFALNSNL 168
D 109 TLWGEFQDGRSARGGCGTNDAAALNASSKEALQIIREKIPKYLFAFALNSNL 168

QY 61 DRIFDKTPEP 70
D 11 DRIFDKTPEP 178
D 169 DRIFDKTPEP 178

RESULT 4
JQ1887
AL1 protein - tomato yellow leaf curl virus (strain Australia)
N:Alternate names: protein
C:Species: tomato yellow leaf curl virus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C:Accession: JQ1887
R:Dry, I.B.; Rigden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.
J. Gen. Virol. 74, 147-151, 1993
A:Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.
A:Reference number: JQ1887; MUID:95139778
A:Accession: JQ1887
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-362 <DR>
A:Cross-references: GH:S53251
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match          61.6%; Score 220; DB 1; Length 362;
Best Local Similarity 52.9%; Pred. No. 1.8e-17;
Matches 45; Conservative 9; Mismatches 15; Indels 16; Gaps 1;

QY 1 TLWGEFQDGRSARGGCGTNDAAALNASSKEALQIIREKIPKYLFAFALNSNL 60
D 1 TLWGEFQDGRSARGGCGTNDAAALNASSKEALQIIREKIPKYLFAFALNSNL 172

```

```

DB 110 TLWGEFQDGRSARGGCGTNDAAALNASSKEALQIIREKIPKYLFAFALNSNL 169

QY 61 DRI-----FDKTPPE 69
D 11 DRI-----FDKTPPE 194
D 170 DRIFTPLLEVYVFLLSSSDVDFPE 194

RESULT 5
S39211
gene C1 protein - tomato yellow leaf curl virus
C:Species: tomato yellow leaf curl virus
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
R:Morris, G.; Hidalgo, E.; Accotto, G.; Moriones, E.
submitted to the EMBL Data Library, August 1993
A:Description: High similarity among the tomato yellow leaf curl virus isolates from
A:Reference number: S39209
A:Accession: S39211
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-352 <DR>
A:Cross-references: ENBL:Z25751; NID:g433655; PIDN:CAA81026.1; PID:g433658
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match          60.8%; Score 217; DB 2; Length 359;
Best Local Similarity 59.4%; Pred. No. 4e-17;
Matches 41; Conservative 9; Mismatches 19; Indels 0; Gaps 0;

QY 2 LWGEFQDGRSARGGCGTNDAAALNASSKEALQIIREKIPKYLFAFALNSNL 61
D 1 LWGEFQDGRSARGGCGTNDAAALNASSKEALQIIREKIPKYLFAFALNSNL 170
D 111 LWGEFQDGRSARGGCGTNDAAALNASSKEALQIIREKIPKYLFAFALNSNL 170

QY 62 DRIFDKTPEP 70
D 11 DRIFDKTPEP 179
D 171 RVFQVPPAP 179

RESULT 6
JQ2327
AL1 protein - Indian cassava mosaic virus
N:Alternate names: replication-associated protein
C:Species: Indian cassava mosaic virus
C:Date: 17-Feb-1994 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C:Accession: JQ2327; S35883
R:Hong, Y.G.; Robinson, D.J.; Harrison, B.D.
J. Gen. Virol. 74, 2437-2443, 1993
A:Title: Nucleotide sequence evidence for the occurrence of three distinct whitefly-
A:Reference number: JQ2326; MUID:94065670
A:Accession: JQ2327
A:Molecule type: DNA
A:Residues: 1-353 <DR>
A:Cross-references: ENBL:Z24758; NID:g395351; PIDN:CAA80891.1; PID:g584046
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match          59.4%; Score 212; DB 2; Length 351;
Best Local Similarity 62.7%; Pred. No. 1.4e-16;
Matches 42; Conservative 6; Mismatches 17; Indels 0; Gaps 0;

QY 4 WGEFQDGRSARGGCGTNDAAALNASSKEALQIIREKIPKYLFAFALNSLDRI 63
D 1 WGEFQDGRSARGGCGTNDAAALNASSKEALQIIREKIPKYLFAFALNSLDRI 172
D 113 WGEFQDGRSARGGCGTNDAAALNASSKEALQIIREKIPKYLFAFALNSLDRI 172

QY 64 FDKTPEP 70
D 11 FDKTPEP 179
D 173 FTKAAPP 179

RESULT 7
S22593
hypothetical protein C4 - tomato yellow leaf curl virus
C:Species: tomato yellow leaf curl virus

```





Db 110 TAWGGEVDIGSARGGQGTANDSYAKALNAGVGSALNLTKEEQPKDYVLNNHNRSL 169  
 QY 61 DRIFDKPTPEP 70  
 Db 170 ERIFAKAPEP 179

## RESULT 12

QJ1870  
 A: protein - tomato mottle virus (isolate Florida)  
 C: Species: tomato mottle virus  
 C: Accession: JQ1870 #sequence\_revision 17-Feb-1994 #text\_change 07-May-1999  
 R: Abouzid, A.W.; Polston, J.E.; Hiebert, E.  
 J. Gen. Virol. 73, 3225-3229, 1992  
 A: Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated from  
 A: Reference number: JQ1869; MUID:93107858  
 A: Accession: JQ1870  
 A: Status: translation not shown  
 A: Residues: 1-358 <R0>  
 A: Cross-References: GB:L14460  
 C: Genetics:  
 A: Map position: segment A  
 C: Superfamily: tomato golden mosaic virus ALL protein

Query Match 56.0%; Score 200; DB 1; Length 358;  
 Best Local Similarity 52.9%; Pred. No. 3,4e-15;  
 Matches 37; Conservative 16; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLWGEFQVDSARGGCGTSDNDAALNASSKEALQIIREKIPEKYLFAFALNSNL 60  
 Db 107 TIEMGEFQIDRSARGGQGSANDSTAKALNAGVGSALVLRKEQPKDFVLNNHNRSL 166  
 QY 61 DRIFDKPTPEP 70  
 Db 167 ERIFAKAPEP 176

## RESULT 13

AL3360  
 A: protein - beet curly top virus  
 C: Species: beet curly top virus  
 C: Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 20-Sep-1999  
 C: Accession: S28360  
 R: Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.  
 EMBO J. 5, 1761-1767, 1986  
 A: Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly to  
 A: Reference number: A40779; MUID:92024070  
 A: Accession: S28360  
 A: Status: translation not shown  
 A: Molecule type: DNA  
 A: Residues: 1-385 <STA>  
 C: Cross-References: GB:M24597; EMBL:X04144; NID:g210678; PIDN:AAA42751.1; PID:g210679  
 C: Superfamily: tomato golden mosaic virus ALL protein

Query Match 56.0%; Score 200; DB 2; Length 385;  
 Best Local Similarity 52.9%; Pred. No. 3,7e-15;  
 Matches 37; Conservative 16; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLWGEFQVDSARGGCGTSDNDAALNASSKEALQIIREKIPEKYLFAFALNSNL 60  
 Db 137 TIEMGEFQIDRSARGGQGSANDSTAKALNAGVGSALVLRKEQPKDFVLNNHNLNA 196  
 QY 61 DRIFDKPTPEP 70  
 Db 197 QKIFQRPDP 206

## RESULT 14

AL3360

S59885  
 replication-associated protein CI - tomato yellow leaf curl virus  
 C: Species: tomato yellow leaf curl virus  
 C: Dates: 14-Jul-1995 #sequence\_revision 01-Mar-1996 #text\_change 20-Sep-1999  
 C: Accession: S59885  
 R: Hong, Y.; Harrison, B.D.  
 submitted to the EMBL Data Library, February 1995  
 A: Description: Nucleotide sequences from tomato leaf curl viruses from different coun  
 d geminiviruses.

A: Reference number: S58346  
 A: Accession: S59885  
 A: Status: preliminary  
 A: Residues: 1-360 <R0>  
 C: Cross-References: EMBL:Z48182; NID:9944838; PIDN:CAA88229.1; PID:g974211  
 C: Superfamily: tomato golden mosaic virus ALL protein

Query Match 54.1%; Score 193; DB 2; Length 360;  
 Best Local Similarity 57.0%; Pred. No. 2,2e-14;  
 Matches 38; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

QY 4 MGFEFQVDSARGGCGTSDNDAALNASSKEALQIIREKIPEKYLFAFALNSLNDRI 63  
 Db 113 FGVFQIDRSARGGQGSANDAYAEATINSKAAALDILREKAPRDVILQPHNLNANDRI 172  
 QY 64 FDKTPPE 69  
 Db 173 FFDPAE 178

## RESULT 15

Q0CVCI  
 A: protein - tomato yellow leaf curl virus  
 A: Alternate names: CI protein  
 C: Species: tomato yellow leaf curl virus  
 C: Dates: 30-Jul-1992 #sequence\_revision 30-Jun-1992 #text\_change 23-Jul-1999  
 C: Accession: D40779  
 R: Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.  
 Virology 185, 151-161, 1991  
 A: Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin  
 A: Reference number: A40779; MUID:92024070  
 A: Accession: D40779  
 A: Status: translation not shown  
 A: Molecule type: DNA  
 A: Residues: 1-357 <NAV>  
 C: Cross-References: GB:X15656; NID:g62204; PIDN:CAA33688.1; PID:g62207  
 C: Superfamily: tomato golden mosaic virus ALL protein

Query Match 53.5%; Score 191; DB 1; Length 357;  
 Best Local Similarity 62.3%; Pred. No. 3,6e-14;  
 Matches 38; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

QY 4 MGFEFQVDSARGGCGTSDNDAALNASSKEALQIIREKIPEKYLFAFALNSLNDRI 63  
 Db 111 FGVSQIDRSARGGQGSANDAYAEATINSKAAALNILEKAPRDVILQPHNLSSLNDRI 170  
 QY 64 F 64  
 Db 171 F 171

Search completed: February 3, 2001, 02:17:40  
 Job time: 2212 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:19:45 : Search time 109.09 Seconds  
(without alignments)  
11.523 Million cell updates/sec

Title: US-09-289-346a-6  
Percent score: 957  
Sequence: 1 TLWGEFQVDSRANGSGQT.....FAFAALNSNLDHLEKTPK 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

1: /cgn2\_6/ptodata/1/isa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/isa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/isa/6.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/isa/PCTUS.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/isa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	210	58.8	359	3 US-08-809-103B-2
2	210	58.8	359	3 US-08-809-103B-4
3	210	58.8	359	3 US-08-809-103B-6
4	210	58.8	359	3 US-08-809-103B-8
5	64.5	18.1	1713	3 US-08-600-982-24
6	64.5	18.1	1713	4 PCT-US94-10261A-24
7	64.5	17.8	2161	1 US-08-452-543A-49
8	63.5	17.8	2161	1 US-08-452-543A-49
9	63.5	17.8	2161	1 US-08-453-543A-51
10	63.5	17.8	2161	2 US-08-223-305C-49
11	63.5	17.8	2161	2 US-08-223-305C-51
12	63.5	17.8	2161	2 US-08-311-363-2
13	58	16.2	446	2 US-08-672-614D-11
14	52	12.4	396	2 US-08-511-485-13
15	52	12.4	396	2 US-08-511-485-13
16	54	15.1	633	1 US-08-458-477A-5
17	54	15.1	633	2 US-08-033-153-5
18	54	15.1	765	2 US-08-663-113-2
19	53.5	15.0	217	1 US-08-621-081A-23
20	53	14.8	266	2 US-07-857-224B-40
21	53	14.8	374	2 US-08-928-692-51
22	52.5	14.9	374	1 US-08-636-080-22
23	52.5	14.7	382	2 US-08-477-451-28
24	52.5	14.7	1781	2 US-08-477-451-11
25	52	14.6	439	1 US-07-762-681A-11
26	52	14.6	439	1 US-07-762-681A-11
27	52	14.6	454	3 US-08-446-100-26
28	52	14.6	454	3 US-08-446-100-27

Sequence 28, Appl  
Sequence 29, Appl  
Sequence 31, Appl  
Sequence 7, Appl  
Sequence 7, Appl  
Sequence 7, Appl  
Sequence 7, Appl  
Sequence 5, Appl  
Sequence 42, Appl  
Sequence 15, Appl  
Sequence 4, Appl  
Sequence 4, Appl  
Sequence 5, Appl  
Sequence 5, Appl  
Sequence 75, Appl

29 52 14.6 454 3 US-08-446-100-28  
30 52 14.6 454 3 US-08-446-100-29  
31 52 14.6 454 3 US-08-446-100-31  
32 52 14.6 482 1 US-07-792-885A-1  
33 52 14.6 482 1 US-08-663-113-2  
34 52 14.6 482 2 US-08-865-477-7  
35 52 14.6 933 3 US-08-293-728-2  
36 52 14.6 4928 3 US-09-036-987A-5  
37 51.5 14.4 373 3 US-08-258-287B-42  
38 51.5 14.4 373 3 US-08-368-704C-42  
39 51.5 14.4 649 3 US-08-911-853-15  
40 51.5 14.4 894 1 US-08-484-352-4  
41 51.5 14.4 894 1 US-08-484-352-4  
42 51 14.3 248 1 US-07-901-701-5  
43 51 14.3 248 1 US-07-988-430-5  
44 51 14.3 248 1 US-08-425-336-5  
45 51 14.3 248 1 US-08-378-761A-75

#### ALIGNMENTS

RESULT 1  
US-08-809-103B-2  
: Sequence 2, Application US/08809103B  
: Patent No. 6133505  
: GENERAL INFORMATION:  
: APPLICANT: GONENBORN, Bruno  
: INVENTOR: GONENBORN, Bruno  
: TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS FOR OBTAINING SAME  
: NUMBER OF SEQUENCES: 17  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: YOUNG & THOMPSON  
: STREET: 745 South 23rd Street  
: CITY: Arlington  
: STATE: Virginia  
: COUNTRY: U.S.A.  
: ZIP: 22202  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patent Release #1.0, Version #1.30  
: CLAIMS: 17  
: APPLICATION NUMBER: US/08/809,103B  
: FILING DATE: 17-MAR-1997  
: CLASSIFICATION: 800  
: PRIOR APPLICATION DATA: FR 94.11040  
: FILING DATE: 15-SEP-1994  
: PRIOR APPLICATION DATA: NO PCT/FR95/01192  
: FILING DATE: 15-SEP-1995  
: ATTORNEY/AGENT INFORMATION:  
: NAME: PATCH, Andrew J.  
: REGISTRATION NUMBER: 32,925  
: REFERENCE/DOCKET NUMBER: US94AL CNR TOM  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (703) 521-2297  
: TELEFAX: (703) 521-2573  
: TELEMAIL: 248423 JPMON  
: INFORMATION FOR SEQ ID NO: 2:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 359 amino acids  
: TYPE: amino acid  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
US-08-809-103B-2

Query Match 58.8%; Score 210; DB 3; Length 359;  
Best Local Similarity 56.5%; Pred. No. 2.3e-20;  
Matches 39; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

QY 2 LWGFEQVQGRSGCGCOTSDNAAEALNASKEEALQIIREKIPKYLFAFALNSNLD 61  
Db 111 LEMGFTQIGRSAGGCGTANDAKAINAGSKQALDVIKELAPDYVLHFRHNSNLD 170  
QY 62 RIFDKTPEP 70  
Db 171 KVFQVPEAP 179

RESULT 2  
US-08-809-103b-4  
Sequence 4, Application US/08809103B  
Patent No. 6133505

GENERAL INFORMATION:  
APPLICANT: GRONENBORN, Bruno  
TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT  
TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
STREET: 745 South 23rd Street  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/809,103B  
FILING DATE: 17-MAR-1997  
PRIORITY INFORMATION:  
PRIORITY NUMBER: FR 94.11040

REGISTRATION NUMBER: 32,925  
REFERENCE/DOCKET NUMBER: USB94AL CNR TOM  
TELEPHONE: (703) 521-2297  
TELEFAX: (703) 685-0573  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 359 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-809-103b-4

Query Match 58.8%; Score 210; Db 3; Length 359;  
Best Local Similarity 56.5%; Pred. No. 2.3e-20;  
Matches 39; Conservative 11; Mismatches 19; Indels 0;  
QY 2 LWGFEQVQGRSGCGCOTSDNAAEALNASKEEALQIIREKIPKYLFAFALNSNLD 61  
Db 111 LEMGFTQIGRSAGGCGTANDAKAINAGSKQALDVIKELAPDYVLHFRHNSNLD 170  
QY 62 RIFDKTPEP 70  
Db 171 KVFQVPEAP 179

RESULT 3  
US-08-809-103b-6

Sequence 6, Application US/08809103B  
Patent No. 6133505  
GENERAL INFORMATION:  
APPLICANT: GRONENBORN, Bruno  
TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT  
TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
STREET: 745 South 23rd Street  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/809,103B  
FILING DATE: 17-MAR-1997  
PRIORITY INFORMATION:  
PRIORITY NUMBER: FR 94.11040  
FILING DATE: 15-SEP-1994  
APPLICATION NUMBER: WO PCT/FR95/01192  
PRIORITY INFORMATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: PATCH, Andrew J.  
REGISTRATION NUMBER: 32,925  
REFERENCE/DOCKET NUMBER: USB94AL CNR TOM

TELEPHONE: (703) 521-2297  
TELEFAX: (703) 685-0573  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 359 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-809-103b-6

Query Match 58.8%; Score 210; Db 3; Length 359;  
Best Local Similarity 56.5%; Pred. No. 2.3e-20;  
Matches 39; Conservative 11; Mismatches 19; Indels 0; Gaps 0;  
QY 2 LWGFEQVQGRSGCGCOTSDNAAEALNASKEEALQIIREKIPKYLFAFALNSNLD 61  
Db 111 LEMGFTQIGRSAGGCGTANDAKAINAGSKQALDVIKELAPDYVLHFRHNSNLD 170  
QY 62 RIFDKTPEP 70  
Db 171 KVFQVPEAP 179

RESULT 4  
US-08-809-103b-8  
Sequence 8, Application US/08809103B  
Patent No. 6133505  
GENERAL INFORMATION:  
APPLICANT: GRONENBORN, Bruno  
TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT  
TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
STREET: 745 South 23rd Street  
CITY: Arlington  
STATE: Virginia







```

1 APPLICATION NUMBER: WO PCT/US89/01406
2 FILING DATE: 04-APR-1989
3 PRIOR APPLICATION DATA:
4 APPLICATION NUMBER: US 07/176,899
5 FILING DATE: 04-APR-1988
6 ATTORNEY/AGENT INFORMATION:
7 NAME: Seidman, Stephanie L.
8 REGISTRATION NUMBER: 3,755
9 REFERENCE/DOCKET NUMBER: 52516 (P519739)
10 TELECOMMUNICATION INFORMATION:
11 TELEPHONE: (619)238-0999
12 TELEFAX: (619)238-0062
13 INFORMATION FOR SEQ ID NO: 51:
14 SEQUENCE CHARACTERISTICS
15 LENGTH: 2161 amino acids
16 TYPE: amino acid
17 STANDARDS: Enghelberg
18 TOPOLOGY: linear
19 MOLECULE TYPE: protein
20 FRAGMENT TYPE: Internal
21 US-08-223-305c-51
22
23 Query Match 17.8%; Score 63.5; DB 2; Length 2161;
24 Base Similarity 28.4%; No. 14;
25 Matches 21; Conservative 5; Mismatches 19; Indels 29; Gaps 31
26
27 QY 4 WGEFVDGSRAGGCGQTS -----NDAAFALNASK 34
28 707 WNAVWDGIMAYGSPSSGIMCVIYFIPLGNYTLNVFLATAVINLADAESLNTAQ 766
29
30 DDB 35 EENQIIRKPIEK 46
31 QY I I I I I
32 DDB 767 EEAERKRIK 760
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
9
```



TELEPHONE: (619)238-0999  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2161 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-311-363-2

Query Match 17.8%; Score 63.5; DB 2; Length 2161;  
Best Local Similarity 28.4%; Pred. No. 14;  
Matches 21; Conservative 5; Mismatches 19; Indels 29; Gaps 1:  
QY 4 WEFQVDSARGQOTS-----NDAASALNASSK 34  
DB 707 WNWVYDGMAYGGSSGMVCIYFILTCGNVILLNVLFAIVNLAQASLNTAQK 766  
OY 35 EFAQLIIRPEK 48  
DB 767 EEAERKKIARK 780

RESULT 13  
US-08-672-814D-11  
Sequence 11592480  
Patent No. 5919912  
GENERAL INFORMATION:  
APPLICANT: Leung, David W.  
APPLICANT: Weeks, Reitha  
TITLE OF INVENTION: MAMMALIAN  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESSES:  
STATE: 201 Elliott Avenue West  
CITY: Seattle  
STATE: Washington  
COUNTRY: U.S.A.  
ZIP: 98119  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" disk, 1.44MB, double side, high density  
OPERATING SYSTEM: Windows NT  
SOFTWARE: WORD 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/672.814D  
FILING DATE: 28-Jun-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Faciszewski, Stephen  
REFERENCE/DOCKET NUMBER: 16, 11  
TELEPHONE: (206)282-7100  
TELEFAX: (206)284-6206  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 446  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: no  
ANTI-SENSE: no  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
STRAIN: Drosophila  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HARLOTYPE:  
TISSUE TYPE:  
CELL TYPE:

CELL LINE:  
FEATURE:  
ANALYSIS: CDP-diacylglycerol synthase (CDS)  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
US-08-672-814D-11

Query Match 16.2%; Score 58; DB 2; Length 446;  
Best Local Similarity 35.3%; Pred. No. 9.3;  
Matches 18; Conservative 10; Mismatches 19; Indels 4; Gaps 2:  
QY 20 TSDAAAEALNASSKEALQIIRKIP-EKYLFAFALNSNLDRIFFKTPE 69  
DB 20 SGSDAANKNSAADSDHVDSEKIPPEKFDVLA---KNLPQGTDKTPE 67

RESULT 14  
US-08-511-485-13  
Sequence 13; Application US/08511485  
Patent No. 5919912  
GENERAL INFORMATION:  
APPLICANT: Korneluk, Robert G.  
APPLICANT: Mackenzie, Alexander E.  
TITLE OF INVENTION: MAMMALIAN IAP GENE FAMILY, PRIMERS,  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/511.485  
FILING DATE: 04-AUG-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 07540/002001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-8900  
TELEFAX: 617/542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 498 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: both  
MOLECULE TYPE: protein  
US-08-511-485-13

Query Match 15.4%; Score 55; DB 2; Length 498;  
Best Local Similarity 26.8%; Pred. No. 28;  
Matches 26; Conservative 8; Mismatches 23; Indels 40; Gaps 4:  
QY 14 ARGCQTSNDAAEALNASSK-----EALQ-----IIRKPE 47  
DB 282 ARGPVYSEVLTATNASSQPATAPTLQADVLMDPAKPAKTLTLDGGVVRNALQR 341  
OY 48 KYL---FAFAALNSNLDRIFD-----KTPEP 70

Db 342 KILSSCAFSTLDLHDFDDGAGAALEVRPEPP 378

```

RESULT 15
US-09-014-969-15
: Sequence 15, Application US/09014969
: Patent No. 5965397
: GENERAL INFORMATION:
: APPLICANT: Jacobs, Kenneth
: APPLICANT: McCoy, John M.
: APPLICANT: Ravallie, Edward R.
: APPLICANT: Reder, James
: APPLICANT: Merber, David
: APPLICANT: Treacy, Maurice
: APPLICANT: Spaulding, Vikki
: APPLICANT: Agostino, Michael J.
: TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
: TITLE OF INVENTION: ENCODING THEM
: NUMBER OF SEQUENCES: 32
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genetics Institute, Inc.
: STREET: 87 Cambridgepark Drive
: CITY: Cambridge
: STATE: MA
: COUNTRY: U.S.A.
: ZIP: 02140
: COMPUTER RELEASABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Sprunger, Suzanne A.
: REGISTRATION NUMBER: 41,323
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 498-9284
: TELEFAX: (617) 876-5651
: INFORMATION PROVIDED:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 335 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-09-014-969-15

```

```

Query Match 15.3%; Score 54.5; DB 2; Length 335;
Best Local Similarity 34.0%; Pred. No. 19;
Matches 16; Conservative 7; Mismatches 19; Indels 5; Gaps 1;

Oy 24 AAARALNASKKEALQITREKIPKYLFAFAALNSMLDRIFDKTPPP 70
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 270 ASLPFVWYDLFEALQVHKVEVITVTYAL-----NIENILTKLWHP 311

```

Search completed: February 3, 2001, 02:19:47  
Job time: 1709 sec

GenCore version 4.5  
 Copyright (c) 1993 - 2000 CompuGen Ltd.  
 OW protein - protein search, using sw model  
 Run on: February 3, 2001, 02:15:28 ; Search time 144.12 Seconds  
 (without alignments)  
 16.608 Million cell updates/sec.

Title: US-09-289-346a-6  
 Perfect score: 357  
 Sequence: 1 TLVMEFQVDRSGRSGQT.....FAFALNSLDRFDKTEP 70

Scoring table: BLOSUM62  
 Gapop 10.0, Gapext 0.5

Searched: 268485 seqs, 34193795 residues  
 Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: A\_Geneseq\_36.\*

1: /SID56/gcgdata/geneseq/genesecp/AA1980.DAT.\*  
 2: /SID56/gcgdata/geneseq/genesecp/AA1981.DAT.\*  
 3: /SID56/gcgdata/geneseq/genesecp/AA1982.DAT.\*  
 4: /SID56/gcgdata/geneseq/genesecp/AA1983.DAT.\*  
 5: /SID56/gcgdata/geneseq/genesecp/AA1984.DAT.\*  
 6: /SID56/gcgdata/geneseq/genesecp/AA1985.DAT.\*  
 7: /SID56/gcgdata/geneseq/genesecp/AA1986.DAT.\*  
 8: /SID56/gcgdata/geneseq/genesecp/AA1987.DAT.\*  
 9: /SID56/gcgdata/geneseq/genesecp/AA1988.DAT.\*  
 10: /SID56/gcgdata/geneseq/genesecp/AA1989.DAT.\*  
 11: /SID56/gcgdata/geneseq/genesecp/AA1990.DAT.\*  
 12: /SID56/gcgdata/geneseq/genesecp/AA1991.DAT.\*  
 13: /SID56/gcgdata/geneseq/genesecp/AA1992.DAT.\*  
 14: /SID56/gcgdata/geneseq/genesecp/AA1993.DAT.\*  
 15: /SID56/gcgdata/geneseq/genesecp/AA1994.DAT.\*  
 16: /SID56/gcgdata/geneseq/genesecp/AA1995.DAT.\*  
 17: /SID56/gcgdata/geneseq/genesecp/AA1996.DAT.\*  
 18: /SID56/gcgdata/geneseq/genesecp/AA1997.DAT.\*  
 19: /SID56/gcgdata/geneseq/genesecp/AA1998.DAT.\*  
 20: /SID56/gcgdata/geneseq/genesecp/AA1999.DAT.\*  
 21: /SID56/gcgdata/geneseq/genesecp/AA2000.DAT.\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	210	58.8	359	17 R88870	Sardinian tomato y
2	210	58.8	359	17 R88871	Sardinian tomato y
3	210	58.8	359	17 R88872	Sardinian tomato y
4	206	57.7	353	18 W34338	Bean golden mosaic
5	206	57.7	353	18 W34332	Bean golden mosaic
6	206	57.7	353	18 W34333	Bean golden mosaic
7	206	57.7	353	18 W34334	Bean golden mosaic
8	206	57.7	353	18 W34335	Bean golden mosaic
9	204	57.1	353	6 P70407	ORF 4 gene product
10	204	57.1	361	18 W34336	Tomato mottle virus
11	204	57.1	361	18 W34324	Tomato mottle virus
12	204	57.1	361	18 W34325	Tomato mottle virus

13	204	57.1	361	18 W34326	Tomato mottle virus
14	197.5	55.3	361	8 P70562	Product of ORF 4 f
15	192	53.8	362	19 W56495	Tobacco leaf curl
16	191	53.5	357	18 W34329	Tomato yellow leaf
17	191	53.5	357	18 W34330	Tomato yellow leaf
18	191	53.5	357	18 W34331	Tomato yellow leaf
19	183	51.3	357	18 W34332	Tomato yellow leaf
20	67	18.8	447	21 Y87089	Human secreted pro
21	67	18.8	447	21 Y66744	Human adult aorta
22	67	18.8	447	21 Y50947	Human secreted pro
23	67	18.8	456	21 Y87190	Human secreted pro
24	64.5	19.1	1713	16 R70148	Human calicivirus Chan
25	63.5	17.8	2161	14 R33545	Human neuronal cal
26	63.5	17.8	2161	14 R33545	Human neuronal cal
27	63.5	17.8	2161	16 R71002	Human neuronal cal
28	63.5	17.8	2161	16 R71002	Human neuronal cal
29	63.5	17.8	2161	19 W63137	Human calcium chan
30	63.5	17.8	2161	19 W63137	Human calcium chan
31	63	17.6	292	21 Y92317	Genitalvirus Rep C1
32	63	17.6	292	21 Y92317	Genitalvirus Rep C1
33	60	16.9	333	18 W34337	Tomato mottle virus
34	59.5	16.7	239	18 W55450	Human ORF 02a
35	59	16.5	318	20 W31895	Soybean cyclin del
36	58.5	16.4	1203	21 Y83275	Candida albicans C
37	58.5	16.4	1203	21 Y84814	Amto acid sequenc
38	57	16.0	62	20 Y11882	Mouse topoisomeras
39	56	15.7	767	20 Y06425	Protein encoded by
40	55.5	15.5	204	18 W9814	Myocardium tube
41	55.5	15.5	266	20 Y21957	Protein encoded by
42	55.5	15.5	266	20 Y21957	Protein encoded by
43	55.5	15.5	299	20 W89829	Antigen encoded by
44	55.5	15.5	299	20 W89849	Antigen from clust
45	55.5	15.5	374	15 R61235	Site specific inte

#### ALIGNMENTS

RESULT 1

R88870 ID R88870 standard; Protein: 359 AA.

XX R88870;

XX R88870;

DT 07-NOV-1996 (first entry)

XX Sardinian tomato yellow leaf curl virus mutated Rep protein (K227A).

XX Nucleotide triphosphate binding site; DNA helicase; RNA helicase;

XX modification; mutation; viral replication; deficient; inhibition;

XX viral resistance; geminivirus; tomato yellow leaf curl virus;

XX Nucleotide triphosphate binding site; DNA helicase; RNA helicase;

XX Nucleotide triphosphate binding site; DNA helicase; RNA helicase;

XX A11 Protein; dominant negative phenotype.

XX Sardinian tomato yellow leaf curl virus.

OS Sardinian tomato yellow leaf curl virus.

XX Key Location/Qualifiers

FT Misc-difference 227 /note- \*wild-type Lys has been replaced by Ala\*

XX W09606573-Al.

XX 21-MAR-1996.

XX 15-SEP-1995; 95WO-FR01192.

XX 15-SEP-1994; 94FR-0011040.

XX (CNRS ) CENT NAT RECH SCI.

XX Gronenborn B;

XX WPI; 1996-179947/18.

DR N-PSDB: T12904.

XX Prodn. of virus-resistant transgenic plants - using mutated genomic

PT sequence from phytopathogenic DNA virus

XX

PS Disclosure: Fig 13; 93pp; French.

XX

CC Mutation of consensus amino acids in the NTP-binding site of

CC geminivirus Rep protein is used to produce replication deficient

CC viruses. The mutated viral nucleic acid is used for producing

CC transgenic plants that are resistant to, or tolerant of, the native

CC virus. The present sequence is a mutant form of the Rep (or C1) virus

CC in which the wild-type Lys227 residue has been changed to an

CC Ala residue; transgenic Nicotiana benthamiana plants generated by

CC transformation with the mutated virus were found to be resistant to

CC STYLGV, i.e. the mutation results in a dominant negative phenotype.

XX

XX Sequence 359 AA:

3Q

Query Match 58.8%; Score 210; DB 17; Length 359;

Best Local Similarity 56.5%; Pred. No. 1e-18;

Matches 39; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

QY 2 LWGFEVDGSRGCGTSDNAAEALMASKEEALQIIREKIPKYLFFAALNSNLD 61

DB 111 lewgftqlgtrsarggqgtandayakalnagksqaldvikelaprdyvhfhainsnld 170

QY 62 RIFDKTPEP 70

DB 171 kvfqvppap 179

RESULT 2

ID R88871 standard; Protein; 359 AA.

XX R88871;

AC R88871;

XX 07-NOV-1996 (first entry)

DE Sardinian tomato yellow leaf curl virus mutated Rep protein (K227H).

XX

XX Nucleotide triphosphate binding site; DNA helicase; RNA helicase;

XX modification; mutation; viral replication; deficient; inhibition;

XX viral resistance; geminivirus; tomato yellow leaf curl virus;

XX Sardinian isolate; STYLGV; transgenic plant; P-loop; C1 protein;

XX A11 protein.

XX Sardinian tomato yellow leaf curl virus.

XX

XX Key Location/Qualifiers

PH Misc-difference 227

FT /note= "wild-type Lys has been replaced by His"

FT

XX W09608573-A1.

XX

XX 21-MAR-1996.

XX

XX 15-SEP-1995; 95WO-FR01192.

XX

XX 15-SEP-1994; 94FR-0011040.

XX

XX (CNRS ) CENT NAT RECH SCT.

XX Gronenborn B;

XX WPI; 1996-179947/18.

XX N-PSDB: T12905.

XX Prodn. of virus-resistant transgenic plants - using mutated genomic

PT sequence from phytopathogenic DNA virus

XX

XX Disclosure: Fig 13; 93pp; French.

XX

CC Mutation of consensus amino acids in the WTP-binding site of

CC geminivirus Rep protein is used to produce replication deficient

CC viruses. The mutated viral nucleic acid is used for producing

CC transgenic plants that are resistant to, or tolerant of, the native

CC virus. The present sequence is a mutant form of the Rep (or C1)

CC protein from the Sardinian isolate of tomato yellow leaf curl virus

CC (STYLGV) in which the wild-type Lys227 residue has been changed to a

CC His residue; transgenic Nicotiana benthamiana plants generated by

CC transformation with the mutated virus were not resistant to STYLGV.

CC However, transgenic plants transformed with a virus in which Lys227 had been

CC replaced by Ala were found to be resistant.

XX

XX Sequence 359 AA:

3Q

Query Match 58.8%; Score 210; DB 17; Length 359;

Best Local Similarity 56.5%; Pred. No. 1e-18;

Matches 39; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

QY 2 LWGFEVDGSRGCGTSDNAAEALMASKEEALQIIREKIPKYLFFAALNSNLD 61

DB 111 lewgftqlgtrsarggqgtandayakalnagksqaldvikelaprdyvhfhainsnld 170

QY 62 RIFDKTPEP 70

DB 171 kvfqvppap 179

RESULT 3

ID R88872 standard; Protein; 359 AA.

XX R88872;

XX R88872;

XX 07-NOV-1996 (first entry)

DE Sardinian tomato yellow leaf curl virus mutated Rep protein (K227R).

XX

XX Nucleotide triphosphate binding site; DNA helicase; RNA helicase;

XX modification; mutation; viral replication; deficient; inhibition;

XX viral resistance; geminivirus; tomato yellow leaf curl virus;

XX Sardinian isolate; STYLGV; transgenic plant; P-loop; C1 protein;

XX A11 protein.

XX Sardinian tomato yellow leaf curl virus.

XX

XX Key Location/Qualifiers

PH Misc-difference 227

FT /note= "wild-type Lys has been replaced by Arg"

FT

XX W09608573-A1.

XX

XX 21-MAR-1996.

XX

XX 15-SEP-1995; 95WO-FR01192.

XX

XX 15-SEP-1994; 94FR-0011040.

XX

XX (CNRS ) CENT NAT RECH SCT.

XX Gronenborn B;

XX WPI; 1996-179947/18.

XX N-PSDB: T12906.

XX Prodn. of virus-resistant transgenic plants - using mutated genomic

PT sequence from phytopathogenic DNA virus

XX

XX Disclosure: Fig 13; 93pp; French.

XX



```

Query Match          57.7%  Score 206;  DB 18;  Length 353;
Best Local Similarity 57.1%  Pred. No. 3.3e-18;
Matches 40;  Conservative 13;  Mismatches 17;  Indels 0;  Gaps 0;

QY 1 TLVWGCFQVDSARGGQTSNDAAAFALNASSKEALQITREKIPKYLFAFALNSWL 60
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 110 TLWQGFQVDSARGGQTSNDAAAFALNASSKEALQITREKIPKYLFAFALNSWL 169
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 DRIFDKTPEP 70
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 170 erlvkvpep 179
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 6
W34333
ID W34333 standard; Protein: 353 AA.
XX
XX W34333;
AC
XX
DT 27-APR-1998 (first entry)
DE Bean golden mosaic geminivirus C1 BGA221 mutant.
XX
XX Geminivirus; BGWV; C1 gene; transdominant mutation;
XX transgenic plant; disease resistance.
XX
XX Bean golden mosaic virus type II isolate Guatemala.
OS
XX
XX W09739110-A1.
PN
XX
PD 23-OCT-1997.
XX
XX 15-APR-1997; 97MO-US06300.
XX
XX 16-APR-1996; 96US-0015517.
XX
XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA
XX (WISC-) WISCONSIN ALUMNI RES FOUND.
XX
XX Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX N-PSDB: T93291.
XX
XX Transgenic plants expressing geminivirus AC1 and C1 wild-type and
XX mutant genes - have increased resistance to geminivirus infection
XX e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
XX golden mosaic geminivirus
XX
XX Example 5; Page 107-109; 132pp; English.
XX
XX This protein comprises a transdominant lethal mutant of the bean
XX golden mosaic virus (BGWV) C1 protein (see W34338) that is required
XX for replication. It is encoded by mutated C1 open reading frame
XX BGA221 (see T93291) and carries a mutation in the NTP-binding
XX domain. The invention involves production of transgenic plants
XX containing DNA encoding production of transgenic plants
XX sequences that negatively interfere in trans with geminiviral
XX replication during infection. Such transgenic plants are resistant
XX to viral infection. The AC1/C1 genes are especially from BGWV,
XX tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
XX and encode polypeptides (see W34324-35) that have mutations in the
XX highly conserved DNA-nicking and/or the NTP-binding domains.
XX
XX Sequence 353 AA:

Query Match          57.7%  Score 206;  DB 18;  Length 353;
Best Local Similarity 57.1%  Pred. No. 3.3e-18;
Matches 40;  Conservative 13;  Mismatches 17;  Indels 0;  Gaps 0;

QY 1 TLVWGCFQVDSARGGQTSNDAAAFALNASSKEALQITREKIPKYLFAFALNSWL 60
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 110 TLWQGFQVDSARGGQTSNDAAAFALNASSKEALQITREKIPKYLFAFALNSWL 169
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 DRIFDKTPEP 70
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 170 erlvkvpep 179
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db 110 TLWQGFQVDSARGGQTSNDAAAFALNASSKEALQITREKIPKYLFAFALNSWL 169
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 DRIFDKTPEP 70
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 170 erlvkvpep 179
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 7
W34334
ID W34334 standard; Protein: 353 AA.
XX
XX W34334;
AC
XX
DT 27-APR-1998 (first entry)
DE Bean golden mosaic geminivirus C1 BGA228 mutant.
XX
XX Geminivirus; BGWV; C1 gene; transdominant mutation;
XX transgenic plant; disease resistance.
XX
XX Bean golden mosaic virus type II isolate Guatemala.
OS
XX
XX W09739110-A1.
PN
XX
PD 23-OCT-1997.
XX
XX 15-APR-1997; 97MO-US06300.
XX
XX 16-APR-1996; 96US-0015517.
XX
XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA
XX (WISC-) WISCONSIN ALUMNI RES FOUND.
XX
XX Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX N-PSDB: T93292.
XX
XX Transgenic plants expressing geminivirus AC1 and C1 wild-type and
XX mutant genes - have increased resistance to geminivirus infection
XX e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
XX golden mosaic geminivirus
XX
XX Example 5; Page 111-112; 132pp; English.
XX
XX This protein comprises a transdominant lethal mutant of the bean
XX golden mosaic virus (BGWV) C1 protein (see W34338) that is required
XX for replication. It is encoded by mutated C1 open reading frame
XX BGA228 (see T93292) and carries a mutation in the NTP-binding
XX domain. The invention involves production of transgenic plants
XX containing DNA encoding production of transgenic plants
XX sequences that negatively interfere in trans with geminiviral
XX replication during infection. Such transgenic plants are resistant
XX to viral infection. The AC1/C1 genes are especially from BGWV,
XX tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
XX and encode polypeptides (see W34324-35) that have mutations in the
XX highly conserved DNA-nicking and/or the NTP-binding domains.
XX
XX Sequence 353 AA:

Query Match          57.7%  Score 206;  DB 18;  Length 353;
Best Local Similarity 57.1%  Pred. No. 3.3e-18;
Matches 40;  Conservative 13;  Mismatches 17;  Indels 0;  Gaps 0;

QY 1 TLVWGCFQVDSARGGQTSNDAAAFALNASSKEALQITREKIPKYLFAFALNSWL 60
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 110 TLWQGFQVDSARGGQTSNDAAAFALNASSKEALQITREKIPKYLFAFALNSWL 169
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 DRIFDKTPEP 70
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 170 erlvkvpep 179
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

RESULT 8
ID W34335 standard; Protein: 353 AA.
XX AC W34335:
XX
XX 27-APR-1998 (first entry)
XX
XX Bean golden mosaic geminivirus CI BGA262 mutant.
XX
XX Geminivirus; BGWV; CI gene; transdominant mutation;
XX transgenic plant; disease resistance.
XX
XX Bean golden mosaic virus type II isolate Guatemala.
XX
XX WO9739110-AL.
XX
XX 23-OCT-1997.
XX
XX 15-APR-1997; 97WO-US06300.
XX
XX 16-APR-1996; 96US-0015517.
XX
XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.
XX (WISC) WISCONSIN ALUMNI RES FOUND.
XX
XX Ahlquist PG, Hanson SF, Lau HT, Maxwell DP, Stout JT;
XX
XX WPI: 1997-526447/48.
XX
XX N-PSDB; 793293.
XX
XX Transgenic plants expressing geminivirus AC1 and CI wild-type and
XX and have increased resistance to geminivirus infection
XX and tomato mottle virus.
XX golden mosaic geminivirus
XX
XX Example 5; Page 115-116; 132pp; English.
XX
XX This protein comprises a transdominant lethal mutant of the bean
XX golden mosaic virus (BGWV) CI protein (see W34338) that is required
XX for the production of transgenic plants.
XX BGAC262 (see T93253) and carries a mutation in the NTP-binding
XX domain. The invention involves production of transgenic plants
XX containing DNA comprising geminivirus CI or AC1 wild-type or mutant
XX sequences that negatively interfere in trans with geminiviral
XX replication during infection. Such transgenic plants are resistant
XX to viral infection. The AC1/CI genes are especially from BGWV,
XX and code polymorphic tomato yellow leaf curl virus (see T93282-93)
XX and code polymorphic tomato yellow leaf curl virus domains in the
XX highly conserved DNA-nicking and/or the NTP-binding domains.
XX
XX Sequence 353 AA:
XX
XX Query Match 57.1%; Score 206; DB 18; Length 353;
XX Best Local Similarity 57.1%; Score 206; DB 18;
XX Matches 40; Conservative 13; Mismatches 17; Indels 0; Gaps 0;
XX
XX Oy 1 TLVWGEFQVDRSGAGCGTSDNDAAEALNASSKEEALQIIRKIPKYLEFAALNSNL 60
XX | : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| : |||
XX Db 110 tiewgfvgdrgsarggqsandsyakalnadsiesaitlkeeqkdvyqlhmrsnl 169
XX
XX Oy 61 DRIFDKTPEP 70
XX | : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| : |||
XX Db 170 erifvktpep 179
XX
XX RESULT 9
XX W34336
XX ID W34336 standard; Protein: 361 AA.
XX AC W34336:
XX
XX 27-APR-1998 (first entry)
XX
XX Tomato mottle virus AC1 protein.
XX
XX Geminivirus; TOMV; AC1 gene; transdominant mutation;
XX transgenic plant; disease resistance.
XX
XX Tomato mottle virus isolate Florida.
XX
XX WO9739110-AL.
XX
XX 23-OCT-1997.
XX
XX 15-APR-1997; 97WO-US06300.
XX
XX 16-APR-1996; 96US-0015517.
XX
XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.
XX (WISC) WISCONSIN ALUMNI RES FOUND.
XX

```

```

AC P70407:
XX
XX 02-MAY-1991 (first entry)
XX
XX ORF 4 gene product of Bean Golden Mosaic virus.
XX
XX Plant vector.
XX
XX Bean golden mosaic virus.
XX
XX JP61257186-A.
XX
XX 14-NOV-1986.
XX
XX 10-MAY-1985; 85JP-0098108.
XX
XX 10-MAY-1985; 85JP-0098108.
XX
XX (TEIJ) TEIJIN KK.
XX
XX WPI: 1987-159662/23.
XX
XX N-PSDB: N70630.
XX
XX New DNA and hybrid DNA - used for recombinant vector of plants.
XX
XX Disclosure; Fig 6; 24pp; Japanese.
XX
XX The sequence encoding this protein may be taken along with the -a
XX DNA and the plant suitable resistance gene, and used to create a
XX recombinant plant vector.
XX
XX See also N70629.
XX
XX Sequence 353 AA:
XX
XX Query Match 57.1%; Score 204; DB 8; Length 353;
XX Best Local Similarity 57.1%; Score 204; DB 8;
XX Matches 40; Conservative 13; Mismatches 17; Indels 0; Gaps 0;
XX
XX Oy 1 TLVWGEFQVDRSGAGCGTSDNDAAEALNASSKEEALQIIRKIPKYLEFAALNSNL 60
XX | : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| : |||
XX Db 110 tiewgfvgdrgsarggqsandsyakalnadsiesaitlkeeqkdvyqlhmrsnl 169
XX
XX Oy 61 DRIFDKTPEP 70
XX | : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| : |||
XX Db 170 erifvktpep 179
XX
XX RESULT 10
XX W34336
XX ID W34336 standard; Protein: 361 AA.
XX AC W34336:
XX
XX 27-APR-1998 (first entry)
XX
XX Tomato mottle virus AC1 protein.
XX
XX Geminivirus; TOMV; AC1 gene; transdominant mutation;
XX transgenic plant; disease resistance.
XX
XX Tomato mottle virus isolate Florida.
XX
XX WO9739110-AL.
XX
XX 23-OCT-1997.
XX
XX 15-APR-1997; 97WO-US06300.
XX
XX 16-APR-1996; 96US-0015517.
XX
XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.
XX (WISC) WISCONSIN ALUMNI RES FOUND.
XX

```





PT	mutant genes - have increased resistance to geminivirus infection
PT	e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT	golden mosaic geminivirus
XX	
XX	Example 3-6; Page 67-69; 132pp; English.
XX	This protein comprises a transmembrane lethal mutant, designated
CC	ToMoV-AC1Δln23, of tomato mottle virus (ToMoV) AC1 protein (see
CC	W34336). It is encoded by a mutant AC1 gene (see T93284) of
CC	tomato mottle virus (ToMoV), and carries 2 mutations in an
CC	NTP-binding domain. The AC1 gene (see also T93284) must be
CC	XPR-pressed for efficient replication of the two genomic components,
CC	DNA-A and DNA-B, of the bipartite ToMoV genome. The invention
CC	comprises production of transgenic plants containing DNA comprising
CC	the sequence of the present invention, which may or may not directly
CC	interfere in trans with geminiviral replication during infection. The
CC	Such transgenic plants are resistant to viral infection. The
CC	AC1/CI genes are especially from ToMoV, tomato yellow leaf curl
CC	virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC	polypeptides (see W34324-35) that have mutations in the highly
CC	conserved DNA-nicking end/or NTP-binding domains.
XX	
XX	Sequence 361 AA;
SQ	
	Query Match 57.1%; Score 204; DB 18; Length 361;
	Best Local Similarity 54.3%; Pred. No. 6.1e-18;
	Matches 38; Conservative 16; Mismatches 16; Indels 0; Gaps
QY	1 TLWRGQVDGHSAGCGTSSNDAAARALNASKKEALQIQIKETPEKYLIFAFANLNSML 60
DB	:          :          :          :          :
QY	61 DRITQRKTEPP 70
DB	:
DB	170 erifakapep 179
RESULT 14	
ID P70562	P70562 standard; Protein: 361 AA.
AC AC	P70562:
DT DT	30-APR-1991 (first entry)
DD DD	Product of ORF 4 from MYMV complementary strand (b).
KX KW	Geminivirus.
OS OS	Mungbean yellow mosaic virus.
PN PN	JF62126982-A.
PF PF	09-JUN-1987.
XX XX	28-NOV-1985; 85JP-0266080.
XX XX	28-NOV-1985; 85JP-0266080.
PA PA	(TEIJ ) TEIJJN KK.
DR DR	N-PSDB; NT0895.
PT PT	Novel DNA and hybrid DNA useful - as vector for recombinant work
PT PT	of plant gene.
XX XX	Disclosure; Fig 8; 21pp; Japanese.
PS PS	The sequence is encoded by ORF 4 which occurs on the complementary
QC QC	strand of the (b) product of the complementary strand of the
	See also J70559-J70567

Search completed: February 3, 2001, 02:15:28  
-Jqb time: 3435 sec

GenCore version 4.5  
 Copyright (c) 1993 - 2000 CompuGen Ltd.  
 OW protein - protein search, using sw model  
 Run on: February 3, 2001, 02:15:28 ; Search time 144.12 Seconds  
 (without alignments)  
 16.608 Million cell updates/sec

Title: US-09-289-346a-7  
 Perfect score: 359  
 Sequence: 1 TLVWGFQVDSARGGQT.....FQFINSLAAIFDKTPEP 70

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5  
 Searched: 268485 seqs, 34193795 residues  
 Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000  
 Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : A.Geneseq.36.\*  
 1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
 2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
 3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
 4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
 5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
 6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
 7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
 8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
 9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
 10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
 11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
 12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
 13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
 14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
 15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
 16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
 17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
 18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
 19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
 20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
 21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Query Score	Match Length	ID	ID	Description
1	213	59.3	353	18	Bean golden mosaic
2	213	59.3	353	18	Bean golden mosaic
3	213	59.3	353	18	Bean golden mosaic
4	213	59.3	353	18	Bean golden mosaic
5	213	59.3	353	18	Bean golden mosaic
6	213	59.3	353	18	Bean golden mosaic
7	211	58.8	359	17	ORF gene product
8	211	58.8	359	17	Sardinian tomato Y
9	211	58.8	359	17	Sardinian tomato Y
10	211	58.8	361	18	Tomato mottle virus
11	211	58.8	361	18	Tomato mottle virus
12	211	58.8	361	18	Tomato mottle virus

13	211	58.8	361	18	W34326
14	208.5	58.1	361	8	P70562
15	194	54.0	362	19	W56495
16	193	53.8	357	18	W34329
17	193	53.8	357	18	W34329
18	193	53.8	357	18	W34331
19	185	51.5	357	18	W34337
20	68	18.9	447	21	V87089
21	68	18.9	447	21	V56744
22	68	18.9	447	21	V50947
23	68	18.9	447	21	V87190
24	63.5	17.7	629	14	K24531
25	63.5	17.7	629	14	K24531
26	63.5	17.7	2161	16	E71001
27	63.5	17.7	2161	16	E71002
28	63.5	17.7	2161	19	W63137
29	63.5	17.7	2161	19	W63149
30	62.5	17.4	512	19	W68473
31	60.5	16.9	131	18	W34327
32	59.5	16.4	335	21	V92318
33	59	16.4	335	21	V92318
34	58.5	16.3	299	18	W55450
35	58	16.2	374	19	W37994
36	58	16.2	374	20	V39875
37	57	15.9	108	16	R82847
38	57	15.9	454	21	V49496
39	57	15.9	454	21	V49497
40	57	15.9	454	21	V49498
41	57	15.9	454	21	V49499
42	57	15.9	454	21	V49500
43	56.5	15.7	188	21	V95043
44	56	15.6	108	16	R82845
45	56	15.6	213	17	W06648

## ALIGNMENTS

RESULT 1  
 W34338  
 ID W34338 standard; Protein: 353 AA.  
 AC W34338;  
 DT 27-APR-1998 (first entry)  
 DE Bean golden mosaic geminivirus C1 protein.  
 XX Geminivirus; BGMV; C1 gene; transdominant mutation;  
 XX transgenic plant; disease resistance.  
 XX Bean golden mosaic virus type II isolate Guatemala.  
 XX W05739110-AL.  
 XX PD 23-OCT-1997.  
 XX 15-APR-1997: 97WO-US06300.  
 XX PR 16-APR-1996: 96US-0015517.  
 XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.  
 XX (WISC) WISCONSIN ALDWINI RES FOUND.  
 XX Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;  
 XX WPI, 1997-526447/48.  
 XX N-PS05; T93914.  
 XX Transgenic plants expressing geminivirus AC1 and C1 wild-type and  
 XX mutant genes - have increased resistance to geminivirus infection  
 XX e.g. tomato mottle virus, tomato yellow leaf curl virus or bean  
 XX golden mosaic geminivirus

XX Example 5; Page 100-102; 132pp; English.

XX This sequence comprises the bean golden mosaic virus (BGMV) CI

XX protein that is required for replication. The invention involves

XX cloning of sequences that negatively interfere in trans

XX wild-type or mutant sequences that negatively interfere in trans

XX with geminiviral replication during infection. Such transgenic

XX plants are resistant to viral infection. The AC1/CI genes are

XX especially from BGMV, tomato mottle virus or tomato yellow leaf

XX curl virus (see T93282-93) and encode polypeptides (see W34324-35)

XX that have mutations in the highly conserved DNA-nicking domain

XX and/or the NTP-binding domains.

XX Sequence 353 AA;

XX

Query Match 59.3%; Score 213; DB 18; Length 353;

Best Local Similarity 58.6%; Pred. No. 1.5e-19;

Matches 41; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLVNGEFGDVGSRAGCGTSDNDAARALNASSKEALQIIRKIPKYLFOFINLSAL 60

DB 110 TLVNGEFGDVGSRAGCGTSDNDAARALNASSKEALQIIRKIPKYLFOFINLSAL 60

QY 61 AAIFDKTPPEP 70

DB 170 erifvkyppep 179

RESULT 2

W34332

ID W34332 standard; Protein: 353 AA.

XX W34332;

DT 27-APR-1998 (first entry)

DE Bean golden mosaic geminivirus CI BGA190 mutant.

XX Geminivirus; BGMV; CI gene; transdominant mutation;

XX transgenic plant; disease resistance.

XX Bean golden mosaic virus type II isolate Guatemala.

PN W09739110-A1.

PD 23-OCT-1997.

PF 15-APR-1997; 97MO-US06300.

PP 16-APR-1996; 96US-0015517.

XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.

PA (WISC) WISCONSIN ALUMNI RES FOUND.

PI Ahlquist PG, Hanson SF, Lau HT, Maxwell DP, Stout JT;

XX WPI: 1997-526447/48.

DR N-PSDB: T93290.

XX Transgenic plants expressing geminivirus AC1 and CI wild-type and

XX mutant genes - have increased resistance to geminivirus infection

XX e.g. tomato mottle virus, tomato yellow leaf curl virus or bean

XX golden mosaic geminivirus

XX Example 5; Page 103-105; 132pp; English.

XX This protein comprises a control mutant of the bean golden mosaic

XX virus (BGMV) CI protein (see W34338) that is required for

XX replication. It is encoded by mutated CI open reading frame

XX BGAC190 (see T93290). The invention involves production of

XX transgenic plants containing geminivirus CI or AC1

CC wild-type or mutant sequences that negatively interfere in trans

CC with geminiviral replication during infection. Such transgenic

CC plants are resistant to viral infection. The AC1/CI genes are

CC especially from BGMV, tomato mottle virus or tomato yellow leaf

CC curl virus (see T93282-93) and encode polypeptides (see W34324-35)

CC that have mutations in the highly conserved DNA-nicking and/or the

CC NTP-binding domains.

XX Sequence 353 AA;

XX

Query Match 59.3%; Score 213; DB 18; Length 353;

Best Local Similarity 58.6%; Pred. No. 1.5e-19;

Matches 41; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLVNGEFGDVGSRAGCGTSDNDAARALNASSKEALQIIRKIPKYLFOFINLSAL 60

DB 110 TLVNGEFGDVGSRAGCGTSDNDAARALNASSKEALQIIRKIPKYLFOFINLSAL 60

QY 61 AAIFDKTPPEP 70

DB 170 erifvkyppep 179

RESULT 3

W34333

ID W34333 standard; Protein: 353 AA.

XX W34333;

DT 27-APR-1998 (first entry)

DE Bean golden mosaic geminivirus CI BGA221 mutant.

XX Geminivirus; BGMV; CI gene; transdominant mutation;

XX transgenic plant; disease resistance.

XX Bean golden mosaic virus type II isolate Guatemala.

PN W09739110-A1.

PD 23-OCT-1997.

PF 15-APR-1997; 97MO-US06300.

PP 16-APR-1996; 96US-0015517.

XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.

PA (WISC) WISCONSIN ALUMNI RES FOUND.

PI Ahlquist PG, Hanson SF, Lau HT, Maxwell DP, Stout JT;

XX WPI: 1997-526447/48.

DR N-PSDB: T93291.

XX Transgenic plants expressing geminivirus AC1 and CI wild-type and

XX mutant genes - have increased resistance to geminivirus infection

XX e.g. tomato mottle virus, tomato yellow leaf curl virus or bean

XX golden mosaic geminivirus

XX Example 5; Page 107-109; 132pp; English.

XX This protein comprises a transdominant lethal mutant of the bean

XX golden mosaic virus (BGMV) CI protein (see W34338) that is required

XX for replication. It is encoded by mutated CI open reading frame

XX BGAC221 (see T93291) and carries a mutation in the start-binding

XX region of the CI gene. The invention involves production of transgenic

XX plants containing DNA comprising geminivirus CI or AC1 wild-type or mutant

XX sequences that negatively interfere in trans with geminiviral

XX replication during infection. Such transgenic plants are resistant

XX to viral infection. The AC1/CI genes are especially from BGMV,

XX tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)

XX and encode polypeptides (see W34324-35) that have mutations in the



```

Oy 61 AAIKDKTEP 70
Db 170 erifkvppep 179

RESULT 6
ID P70407 standard; Protein; 353 AA.
XX
XX P70407;
XX
XX P70407;
XX
Dt 02-WAY-1991 (first entry)
XX
XX ORF 4 gene product of Bean Golden Mosaic virus.
DE
XX Plant vector.
XX
XX Bean golden mosaic virus.
XX
XX JF61257186-A.
XX
XX 14-NOV-1986.
XX
XX 10-WAY-1985; 85JP-0098108.
XX
XX 10-WAY-1985; 85JP-0098108.
XX
XX (TEIJ ) TEIJIN KK.
XX
XX WPI; 1987-159662/23.
XX
XX New DNA and hybrid DNA - used for recombinant vector of plants.
XX
XX Disclosure; Fig 6; 24pp; Japanese.
XX
XX The sequence encoding this protein may be taken along with the -a
CC DNA sequence and a suitable resistance gene, and used to create a
CC recombinant plant vector.
XX
XX See also N70629.
XX
XX Sequence 353 AA:

Query Match 58.8%; Score 211; DB 8; Length 353;
Best Local Similarity 58.6%; Pred. No. 2.7e-19;
Matches 41; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

Oy 1 PLWGFQVDCRSARGCGTSDNAAALNASSKEALQIIRKPEKYLQFHNLSAL 60
Db 110 tlewgfvqdgrrsggqgsamsyakaInmdslessdtlleqphprlqimrlsml 169

Oy 61 AAIKDKTEP 70
Db 170 erifkvppep 179

RESULT 7
ID R88870 standard; Protein; 359 AA.
XX
XX R88870;
XX
XX 07-NOV-1996 (first entry)
XX
XX Sardinian tomato yellow leaf curl virus mutated Rep protein (K227A).
XX
XX Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
XX modification; mutation; viral replication; deficient; inhibition;
XX viral resistance; geminivirus; tomato yellow leaf curl virus;
XX Sardinian isolate; STYICV; transgenic plant; P-loop; CI protein;
XX All protein; dominant negative phenotype.
XX
XX All protein type.
XX
XX Key Location/Qualifiers

```

```

XX OS Sardinian tomato yellow leaf curl virus.
XX
XX Key Location/Qualifiers
XX
XX Misc-difference 22;
XX /Note= "wild-type Lys has been replaced by Ala"
XX
XX W09608573-A1.
XX
XX 21-MAR-1996.
XX
XX 15-SEP-1995; 95WO-EN01192.
XX
XX 15-SEP-1994; 94FR-0011040.
XX
XX (CNRS ) CENT NAT RECH SCI.
XX
XX Gronenborn B;
XX
XX WPI. 1996-176947/16.
XX
XX N-PSDB; T12904.
XX
XX Prodn. of virus-resistant transgenic plants - using mutated genomic
XX sequence from phytopathogenic DNA virus
XX
XX Disclosure; Fig 13; 93pp; French.
XX
XX Mutation of consensus amino acids in the WIP-binding site of
XX geminivirus Rep protein used to produce tobacco etch virus deficient
XX viruses. The mutated viral nucleic acid is used for producing
XX transgenic plants that are resistant to, or tolerant of, the native
XX virus. The present sequence is a mutant form of the Rep (or CI)
XX protein from the Sardinian isolate of tomato yellow leaf curl virus
XX in which the wild-type Lys227 residue has been changed to an
XX Ala residue; transgenic Nicotiana benthamiana plants generated by
XX transformation with the mutated virus were found to be resistant to
XX STYICV, i.e. the mutation results in a dominant negative phenotype.
XX
XX Sequence 359 AA:

Query Match 58.8%; Score 211; DB 17; Length 359;
Best Local Similarity 56.5%; Pred. No. 2.8e-19;
Matches 39; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

Oy 2 PLWGFQVDCRSARGCGTSDNAAALNASSKEALQIIRKPEKYLQFHNLSALA 61
Db 111 tlewgfvqdgrrsggqgsamsyakaInadgsqaldvikelaprdyvihfmlnsld 170

Oy 62 AIFDKTPEP 70
Db 171 kvfqppep 179

RESULT 8
ID R88871 standard; Protein; 359 AA.
XX
XX R88871;
XX
XX 07-NOV-1996 (first entry)
XX
XX Sardinian tomato yellow leaf curl virus mutated Rep protein (K227H).
XX
XX Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
XX modification; mutation; viral replication; deficient; inhibition;
XX viral resistance; geminivirus; tomato yellow leaf curl virus;
XX Sardinian isolate; STYICV; transgenic plant; P-loop; CI protein;
XX All protein.
XX
XX Sardinian tomato yellow leaf curl virus.
XX
XX Key Location/Qualifiers

```

```

FT  Misc-difference 227
XX  /note= *wild-type Lys has been replaced by His*
XX  WO9608573-AL.
XX  21-MAR-1996.
XX  15-SEP-1995; 95WO-FR01192.
XX  15-SEP-1994; 94FR-0011040.
XX  (CNRS ) CENT NAT RECH SCI.
XX  Gronenborn B;
XX  WPI: 1996-179947/18.
XX  N-PSDB; T12905.
XX  Prodn. of virus-resistant transgenic plants - using mutated genomic
XX  sequence from phytopathogenic DNA virus
XX  Disclosure; Fig 13; 93pp; French.
XX  Mutation of consensus amino acids in the NTP-binding site of
XX  geminivirus Rep protein is used to produce replication deficient
XX  viruses. The mutated viral nucleic acid is used for producing
XX  transgenic plants that are resistant to, or tolerant of, the native
XX  virus. The present sequence is a mutant form of the Rep (or C1)
XX  protein from the Sardinian isolate of tomato yellow leaf curl virus
XX  (STYLVCV) in which the wild-type Lys227 residue has been changed to a
XX  His residue; transgenic Nicotiana benthamiana plants generated by
XX  transformation with the mutated virus were not resistant to STYLVCV.
XX  In contrast, plants transformed with a virus in which Lys227 had been
XX  replaced by Ala were found to be resistant.
XX  Sequence 359 AA;
SQ

Query Match 58.8%; Score 211; DB 17; Length 359;
Best Local Similarity 56.5%; Pred. No. 2.8e-19;
Matches 39; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

Qy 2 LVWGEFQVDSRGSGCGTSNDAAALNASKSEALQITREKIPKYLQFHNLSALA 61
Db 111 lewgtfqldgrsrggqatandayakainogaksgaldvikelaprdyvhfhinsold 170

Qy 62 AFDKTEPE 70
Db 171 kvfgvppap 179

RESULT 9
R88872 ID R88872 standard; Protein; 359 AA.
XX AC R88872;
XX 07-NOV-1996 (first entry)
XX Sardinian tomato yellow leaf curl virus mutated Rep protein (K227R).
DE Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW modification; mutation; viral replication; deficient; inhibition;
KW viral resistance; geminivirus; tomato yellow leaf curl virus;
KW Sardinian isolate; STYLVCV; transgenic plant; p-loop; C1 protein;
XX All protein.
XX Sardinian tomato yellow leaf curl virus.
XX Key Location/Qualifiers
XX Misc-difference 227
FT /note= *wild-type Lys has been replaced by Arg*
XX

```

```

PN WO9608573-AL.
XX 21-MAR-1996.
XX 15-SEP-1995; 95WO-FR01192.
XX 15-SEP-1994; 94FR-0011040.
XX (CNRS ) CENT NAT RECH SCI.
XX Gronenborn B;
XX WPI: 1996-179947/18.
XX N-PSDB; T12905.
XX Prodn. of virus-resistant transgenic plants - using mutated genomic
XX sequence from phytopathogenic DNA virus
XX Disclosure; Fig 13; 93pp; French.
XX Mutation of consensus amino acids in the NTP-binding site of
XX geminivirus Rep protein is used to produce replication deficient
XX viruses. The mutated viral nucleic acid is used for producing
XX transgenic plants that are resistant to, or tolerant of, the native
XX virus. The present sequence is a mutant form of the Rep (or C1)
XX protein from the Sardinian isolate of tomato yellow leaf curl virus
XX (STYLVCV) in which the wild-type Lys227 residue has been changed to
XX an Arg residue; transgenic Nicotiana benthamiana plants generated by
XX transformation with the mutated virus were not resistant to STYLVCV.
XX In contrast, plants transformed with a virus in which Lys227 had been
XX replaced by Ala were found to be resistant.
XX Sequence 359 AA;
SQ

Query Match 58.8%; Score 211; DB 17; Length 359;
Best Local Similarity 56.5%; Pred. No. 2.8e-19;
Matches 39; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

Qy 2 LVWGEFQVDSRGSGCGTSNDAAALNASKSEALQITREKIPKYLQFHNLSALA 61
Db 111 lewgtfqldgrsrggqatandayakainogaksgaldvikelaprdyvhfhinsold 170

Qy 62 AFDKTEPE 70
Db 171 kvfgvppap 179

RESULT 10
W34336 ID W34336 standard; Protein; 361 AA.
XX AC W34336;
XX 27-APR-1998 (first entry)
XX Tomato mottle virus AC1 protein.
XX Geminivirus; ToMoV; AC1 gene; transdominant mutation;
XX transgenic plant; disease resistance.
XX Tomato mottle virus isolate Florida.
OS WO9739110-AL.
XX 23-OCT-1997.
XX 15-APR-1997; 97WO-US06300.
XX 16-APR-1996; 96US-0015517.
XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.
XX (WISC ) WISCONSIN ALUMNI RES FOUND.
XX

```

XX PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;  
 XX WP1: 1997-526447/48.  
 XX N-PSDB: T93284.  
 XX Transgenic plants expressing geminivirus AC1 and CI wild-type and  
 PT mutant genes - have increased resistance to geminivirus infection  
 PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean  
 PT golden mosaic geminivirus  
 XX Example 3.3: Page 57-58: 132pp: English.  
 XX This protein comprises the wild-type AC1 protein of tomato mottle  
 CC virus (ToMoV), a geminivirus that has a bipartite genome. The AC1  
 CC gene (see T93294) must be expressed for efficient replication of  
 CC the two genomic components, DNA-A and DNA-B. The AC1 protein has a  
 CC DNA binding site specific to the DNA-A common region, a DNA nicking  
 CC activity, and an NTP-binding activity. The invention involves  
 CC production of transgenic plants containing DNA comprising AC1 or CI  
 CC genes, and methods for using such plants. Such transgenic  
 CC plants are resistant to viral infection. The AC1/CI genes are  
 CC especially from ToMoV, tomato yellow leaf curl virus or bean golden  
 CC mosaic geminivirus (see T93282-93) and encode polypeptides (see  
 CC W34324-35) that have mutations in the highly conserved DNA-nicking  
 CC domain and/or the NTP-binding domain.  
 XX Sequence 361 AA;

Query Match 58.8%; Score 211; DB 18; Length 361;  
 Best Local Similarity 55.7%; Pred. No. 2.8e-19;  
 Matches 39; Conservative 15; Mismatches 16; Indels 0; Gaps 0;  
 QY 1 TLVKGFEQVDGRSGCGQTSNDAAAPALNASSKEALQITREKIPKYLQFHNLSAL 60  
 Db 110 tlewgdfqldgrsrggqsandsyakaInassvqsalavireeqkdfvlgqhmirsml 169  
 QY 61 AAIFDKTPEP 70  
 Db 170 erifakepep 179

RESULT 11  
 W34324  
 ID W34324 standard; Protein: 361 AA.  
 AC W34324;  
 DT 27-APR-1998 (first entry)  
 XX Tomato mottle virus AC1 mutant ToMoV-AC1dml.  
 DE Geminivirus; ToMoV-AC1dml: AC1 gene; transdominant mutation;  
 KW transgenic plant; disease resistance.  
 XX Tomato mottle virus isolate Florida.  
 OS Synthetic.  
 XX W09739110-A1.  
 PN W09739110-A1.  
 XX 23-OCT-1997.  
 PD 23-OCT-1997.  
 XX 15-APR-1997; 97WO-US06300.  
 XX 16-APR-1996; 96US-0015517.  
 PR (SEMI-) SEMINIS VEGETABLE SEEDS INC.  
 XX (WISC ) WISCONSIN ALUMNI RES FOUND.  
 PA Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;  
 XX WP1: 1997-526447/48.  
 XX N-PSDB: T93283.

DR WP1: 1997-526447/48.  
 DR N-PSDB: T93282.  
 XX Transgenic plants expressing geminivirus AC1 and CI wild-type and  
 PT mutant genes - have increased resistance to geminivirus infection  
 PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean  
 PT golden mosaic geminivirus  
 XX Example 3.4: Page 60-62: 132pp: English.  
 XX This protein comprises a transdominant lethal mutant, designated  
 CC ToMoV-AC1dml, of tomato mottle virus (ToMoV) AC1 protein (see  
 CC T93294) that has a bipartite genome. The AC1 gene (see also  
 CC T93294) must be expressed for efficient replication of the two  
 CC NTP-binding domains. The AC1 gene (see also T93294) must be  
 CC expressed for efficient replication of the two genomic components,  
 CC DNA-A and DNA-B, of the bipartite ToMoV genome. The invention  
 CC involves production of transgenic plants containing DNA comprising  
 CC geminivirus AC1 or CI wild-type or mutant sequences that negatively  
 CC interfere in trans with geminiviral replication during infection.  
 CC Such transgenic plants are resistant to viral infection. Such  
 CC AC1/CI genes are especially from ToMoV, tomato yellow leaf curl  
 CC virus or bean golden mosaic geminivirus (see T93282-93) and encode  
 CC polypeptides (see W34324-35) that have mutations in the highly  
 CC conserved DNA-nicking and/or NTP-binding domains.  
 XX Sequence 361 AA;

Query Match 58.8%; Score 211; DB 18; Length 361;  
 Best Local Similarity 55.7%; Pred. No. 2.8e-19;  
 Matches 39; Conservative 15; Mismatches 16; Indels 0; Gaps 0;  
 QY 1 TLVKGFEQVDGRSGCGQTSNDAAAPALNASSKEALQITREKIPKYLQFHNLSAL 60  
 Db 110 tlewgdfqldgrsrggqsandsyakaInassvqsalavireeqkdfvlgqhmirsml 169  
 QY 61 AAIFDKTPEP 70  
 Db 170 erifakepep 179

RESULT 12  
 W34325  
 ID W34325 standard; Protein: 361 AA.  
 AC W34325;  
 DT 27-APR-1998 (first entry)  
 XX Tomato mottle virus AC1 mutant ToMoV-AC1dml.  
 XX Geminivirus; ToMoV-AC1dml: AC1 gene; transdominant mutation;  
 KW transgenic plant; disease resistance.  
 XX Tomato mottle virus isolate Florida.  
 OS Synthetic.  
 XX W09739110-A1.  
 PN W09739110-A1.  
 XX 23-OCT-1997.  
 PD 23-OCT-1997.  
 XX 15-APR-1997; 97WO-US06300.  
 XX 16-APR-1996; 96US-0015517.  
 PR (SEMI-) SEMINIS VEGETABLE SEEDS INC.  
 XX (WISC ) WISCONSIN ALUMNI RES FOUND.  
 PA Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;  
 XX WP1: 1997-526447/48.  
 XX N-PSDB: T93283.



PT		tant genes - have increased resistance to geminivirus infection
e.g.		tomato mottle virus, tomato yellow leaf curl virus or bean
PT		golden mosaic geminivirus
XX		
Example 3.6:	Page 67-69; 132pp; English.	
XX		
CC	This protein comprises a transdominant lethal mutant, designated	
CC	TACidit <sub>1</sub> , cloned from a mottled leaflet (Tomov) ACI gene (see	
CC	W34316), which encodes a putative ACI gene product (see T9328).	
CC	The ACI gene has been shown to be involved in replication of	
CC	tomato yellow virus (Tomov), and carries 2 mutations in an	
CC	NTP-binding domain. The ACI gene (see also T93294) must be	
CC	expressed for efficient replication of the two genomic components,	
CC	DNA-A and DNA-B, of the bipartite Tomov genome. The invention	
CC	involves production of transgenic plants containing DNA comprising	
CC	geminivirus ACI or CI wild-type or mutant sequences that negatively	
CC	suppresses in planta geminivirus replication during infection.	
CC	Such transgenic plants are suitable as hosts for cloning ACI/CI	
CC	genes, especially from Tomov, tomato yellow leaf curl	
CC	virus or bean golden mosaic geminivirus (see T93282-93) and encode	
CC	polypeptides (see J4324-35) that have mutations in the highly	
CC	conserved DNA-nicking and/or NTP-binding domains.	
XX		
Sequence	361 AA;	
XX		
Query Match	58.8%; Score 211; DB 18; Length 361;	
Best Local Similarity	55.7%; Pred. No. 2,8e-19;	
Matches	39; Conservative 15; Mismatches 16; Indels 0; Gaps	
Qy	1 TLWGSEFVGVSARGCGTSDNDAAALNASKKEEALQITREPKYLFQFHMSAL 60	
DB	110 LLEGGDTGDSRGYQGSGSNDYSKADINASSVGLAVIRCEPKDVFHMAIRSL 169	
Oy	61 AAFDKTPPEP 70	
DB	170 erifakapep 179	
RESULT 14		
P70562	ID TO P70562 Standard; Protein; 361 AA.	
AC	P70562;	
AC	30-APR-1991 (First entry)	
DT	Product of ORF 4 from MYMV complementary strand (b).	
XX	Geminivirus.	
XX	Mungbean yellow mosaic virus.	
OS	JPG2126982-A.	
PN	09-JUN-1987.	
PF	28-NOV-1985; 85JP-02666080.	
XX	28-NOV-1985; 85JP-02666080.	
XX	(TEIJ ) TELJIN KK.	
FA	WPI; 1987-196308/28.	
DR	N-PSIB; W70895.	
PT	Novel DNA and hybrid DNA useful - as vector for recombinant work	
PT	of plant gene.	
XX	Disclosure; Fig 8; 21pp; Japanese.	
XX	The sequence is encoded by ORF 4 which occurs on the complementary	
XX	strand of the (b) component of the geminivirus.	
CC	See AJ380-P70559-P70567.	

Search completed: February 3, 2001, 02:15:29  
Job time: 3436 sec

Result	Query No.	Score	Match	Length	DB	ID	Description
	1	281	77.2	226	12	009727	leonus mo
	2	281	77.2	226	12	09WHF6	Q9WHF6 tomato mild
	3	281	77.2	361	12	Q67574	Q67574 bean golden
	4	279	76.1	225	12	09QDB1	Q9QDB1 cowpea gold
	5	277	76.1	364	12	09QDB5	Q9QDB5 sweet potato
	6	274	75.3	185	12	Q86593	Q86593 sida golden
	7	274	75.3	185	12	Q86593	Q86593 sida golden
	8	266	73.1	233	12	08YLAA	Q8YLAA macropittium
	9	261	71.7	190	12	09Z089	Q9Z089 tobacco lea
	10	261	71.7	190	12	Q9Z084	Q9Z084 tobacco lea
	11	260	71.4	190	12	09W827	Q9W827 tobacco lea
	12	260	71.4	208	12	Q9Z0C4	Q9Z0C4 tobacco lea
	13	260	71.4	234	12	Q93180	Q93180 potato yell
	14	258	70.9	208	12	Q9Z0C0	Q9Z0C0 tobacco lea
	15	258	70.9	208	12	Q9Z0B8	Q9Z0B8 tobacco lea
	16	257	70.3	203	12	Q9Z086	Q9Z086 tobacco lea
	17	225	69.1	203	12	Q9Z083	Q9Z083 tobacco lea
	18	251	69.0	363	12	Q73577	Q73577 cotton leaf
	19	251	69.0	363	12	Q73577	Q73577 cotton leaf
	20	250	68.7	363	12	Q72705	Q72705 cotton leaf



RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;  
 RT "Detection of a geminivirus infecting sweet potato in the United  
 States."; 82:1253-1257(1998).  
 DR EMBL: AF040436:AA047177.1; -;  
 DR PFAM: PF00799; Gemini\_1; 1.  
 DR INTERPRO: IPR001301; -;  
 DR PFAM: PF00799; Gemini\_1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATLVL1.  
 SQ SEQUENCE 364 AA; 40680 MW; 5F79752431A906E CRC64;

Query Match 76.1%; Score 277; DB 12; Length 364;  
 Best Local Similarity 79.4%; Pred. No. 2.9e-22;  
 Matches 54; Conservative 3; Mismatches 11; Indels 0; Gaps 0;  
 Oy 1 TUVGSEFVQDSARGCQTSNOLLLEALNASSKEALQIIREKIPKYLQFHNLSNL 60  
 Db 110 TIVGSEFVQDSARGCQTSNOLLLEALNASSKEALQIIREKIPKYLQFHNLSNL 169  
 Oy 61 DRIFDKTP 68  
 Db 170 DRIFSEPP 177

RESULT 6  
 ID Q98693 PRELIMINARY; PRT; 185 AA.  
 AC Q98693;  
 DT 01-FEB-1997 (TRENBLrel. 02, Created)  
 DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)  
 DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)  
 DE REP PROTEIN (FRAGMENT).  
 OS Sida golden mosaic virus  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=51034;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAMAICA;  
 RA Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;  
 DR EMBL: AF098940:AA017850.1; -;  
 DR PFAM: PF00799; Gemini\_1; 1.  
 DR INTERPRO: IPR001301; -;  
 DR PFAM: PF00799; Gemini\_1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATLVL1.  
 FT NON\_TER 185 181  
 FT NON\_TER 185 181  
 SQ SEQUENCE 185 AA; 20975 MW; 3913850A025A5EE1 CRC64;

Query Match 75.3%; Score 274; DB 12; Length 185;  
 Best Local Similarity 70.0%; Pred. No. 2.9e-22;  
 Matches 49; Conservative 12; Mismatches 9; Indels 0; Gaps 0;  
 Oy 1 TUVGSEFVQDSARGCQTSNOLLLEALNASSKEALQIIREKIPKYLQFHNLSNL 60  
 Db 89 TIENGVOIDRSARGCQTSNOLLLEALNASSKEALQIIREKIPKYLQFHNLSNL 148  
 Oy 61 DRIFDKTP 70  
 Db 149 DRIFSKPP 158

RESULT 7  
 ID P88975 PRELIMINARY; PRT; 149 AA.  
 AC P88975;  
 DT 01-MAY-1997 (TRENBLrel. 03, Created)  
 DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)

DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)  
 DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).  
 GN ACL.  
 OS Macropitium golden mosaic geminivirus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=51076;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAMAICA;  
 RA Roye M.E.;  
 DR EMBL: 075278; AA036919.1; -;  
 DR PFAM: PF00799; Gemini\_1; 1.  
 DR INTERPRO: IPR001301; -;  
 DR PFAM: PF00799; Gemini\_1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATLVL1.  
 FT NON\_TER 149 149  
 FT NON\_TER 149 149  
 SQ SEQUENCE 149 AA; 16785 MW; 24CF5ED049CD508 CRC64;

Query Match 73.1%; Score 266; DB 12; Length 149;  
 Best Local Similarity 68.6%; Pred. No. 1.7e-21;  
 Matches 48; Conservative 12; Mismatches 10; Indels 0; Gaps 0;  
 Oy 1 TUVGSEFVQDSARGCQTSNOLLLEALNASSKEALQIIREKIPKYLQFHNLSNL 60  
 Db 52 TIENGVOIDRSARGCQTSNOLLLEALNASSKEALQIIREKIPKYLQFHNLSNL 111  
 Oy 61 DRIFDKTP 70  
 Db 112 DRIFMDPEP 121

RESULT 8  
 ID Q9YL44 PRELIMINARY; PRT; 233 AA.  
 AC Q9YL44;  
 DT 01-MAY-1999 (TRENBLrel. 10, Created)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
 DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)  
 DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).  
 OS Macropitium golden mosaic geminivirus  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=51076;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAMAICA STRAIN 1;  
 RA Roye M.E.;  
 DR EMBL: AF098940:AA017850.1; -;  
 DR PFAM: PF00799; Gemini\_1; 1.  
 DR INTERPRO: IPR001301; -;  
 DR PFAM: PF00799; Gemini\_1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATLVL1.  
 FT NON\_TER 233 233  
 FT NON\_TER 233 233  
 SQ SEQUENCE 233 AA; 26355 MW; AA490AF4D2166A02 CRC64;

Query Match 73.1%; Score 266; DB 12; Length 233;  
 Best Local Similarity 68.6%; Pred. No. 2.8e-21;  
 Matches 48; Conservative 12; Mismatches 10; Indels 0; Gaps 0;  
 Oy 1 TUVGSEFVQDSARGCQTSNOLLLEALNASSKEALQIIREKIPKYLQFHNLSNL 60  
 Db 52 TIENGVOIDRSARGCQTSNOLLLEALNASSKEALQIIREKIPKYLQFHNLSNL 111  
 Oy 61 DRIFDKTP 70  
 Db 112 DRIFMDPEP 121

```

Query Match          71.7%  Score 261.1  DB 12:  Length 190;
Best Local Similarity 60.0%  Pred. No. 7,7e-21;
Matches 51; Conservative 8; Mismatches 10; Indels 16; Gaps 1;

OY 1 TLWGEQVQDGRGARGCQTSDNLLLEALNASKEEAQIIRIKPEKLYQFHNINSNL 60
DB 95 TLWGTFTQDGRGARGQNADCAEALNASKAELAIIRIKPLKDFQFHNINSNL 144
OY -----FDTKE 69
DB 111
OY 145 DRIFAPLEVFCVPTFASDFQVPE 169
DB 111

RESULT 10
Q20806 PRELIMINARY; PRT: 190 AA.
AC Q92084
DC Q92084: PRELIMINARY;
DT 01-MAY-1999 (TRENDELrel. 10, Created)
DD 01-MAY-1999 (TRENDELrel. 10, Last sequence update)
DT 01-MAY-2000 (TRENDELrel. 13, Last annotation update)
DE C1 AND C4 GENES, CLONE YOKOHAMA5-2,
DE PARTIAL AND COMPLETE CDS (FRAGMENT).
DE Tobacco leaf curl virus.
OS Viruses; ssDNA viruses, Geminiviridae; Begomovirus.
ON NCBI_TaxID=57762;
OX [1]
RC STRAIN=YOKOHAMA5;
RC K. Onishi S. Ishii I. Yahara T.;
RC M. K. Onishi S. Yahara T.
RL J-Plant Res. 110:247-257(1997).

```

DR	EMBL:	AB001318;	BAA34039.1; -
DR	INTERPRO:	IPR001191;	-
DR	INTERPRO:	IPR001301;	-
DR	PFAM:	PF00799;	Gemini_ALI; 1.
DR	PRINTS:	PR00227;	GEMCONTALI.
DR	PRINTS:	PR00228;	GEMCOTCLVLI.
FT	NON_TER	1	
FT	NON_TER	190	
SQ	SEQUENCE	190 AA; 21444 MW; AACLC2943EFJOAD CRC64;	
Query Match            71.4%; Score 261; DB 12; Length 190;			
Best local similarity   60.0%; Pred.No. 7.7e-21;			
Matches 51; Conservative 8; Mismatches 10; Indels 16; Gaps 1;			
Oy	1	TLMWGTFQVDSRGAGCGQTSDNLLLEALNAASKEEALQITREKLPKYLFQPHINLSNL 60	
Dy	1	TLMWGTFQVDSRGAGCGQTSDNLLLEALNAASKEEALQITREKLPKYLFQPHINLSNL 60	
Dd	85	TLMWGTQIDGSRAGGCONANDCAEALNAASKAEALATIREKLPGDPTFOYHINSNL 144	
Oy	61	DRI-----FDKTP 69	
Dy	111		
Dd	145	DRIAPPLEVFCVFSSDFQVPE 169	
Oy	111		
Dy	111		
Dd	145	DRIAPPLEVFCVFSSDFQVPE 169	
RESULT	11		
ID	Q9M827	PRELIMINARY; PRT; 190 AA.	
AC	Q9M827;		
DT	01-NOV-1999	(TrEMBLrel. 12. Created)	
DT	01-NOV-1999	(TrEMBLrel. 12. Last sequence update)	
DT	01-MAY-2000	(TrEMBLrel. 13. Last annotation update)	
DE	C1	PROTEIN [FRAGMENT].	
OS	Tobacco leaf curl virus.		
QS	Viruses; ssRNA viruses; Geminiviridae; Begomovirus.		
OX	NCBI_Taxid=67762;		
RN	[1]	SEQUENCE FROM N.A.	
RC	SHRYAN-CORAL; S. Ichihara T.		
RL	Molecular phylogeny of geminivirus infecting wild plants in Japan."		
RL	J. Plant Res. 110:247-257(1997).		
EMBL:	AB001303; BAA34010.1; -		
DR	INTERPRO:	IPR001191; -	
DR	INTERPRO:	IPR001301; -	
DR	PFAM:	PF00799; Gemini_ALI; 1.	
DR	PRINTS:	PR00227; GEMCONTALI.	
DR	PRINTS:	PR00228; GEMCOTCLVLI.	
FT	NON_TER	1	
FT	NON_TER	190	
SQ	SEQUENCE	190 AA; 21444 MW; 93C3742ABEDB7EB CRC64;	
Query Match            71.4%; Score 260; DB 12; Length 190;			
Best local similarity   60.0%; Pred.No. 9.9e-21;			
Matches 51; Conservative 8; Mismatches 10; Indels 16; Gaps 1;			
Oy	1	TLMWGTFQVDSRGAGCGQTSDNLLLEALNAASKEEALQITREKLPKYLFQPHINLSNL 60	
Dy	1	TLMWGTFQVDSRGAGCGQTSDNLLLEALNAASKEEALQITREKLPKYLFQPHINLSNL 60	
Dd	85	TLMWGTQIDGSRAGGCONANDCAEALNAASKAEALATIREKLPGDPTFOYHINSNL 144	
Oy	61	DRI-----FDKTP 69	
Dy	111		
Dd	145	DRIAPPLEVFCVFSSDFQVPE 169	
Oy	111		
Dy	111		
Dd	145	DRIAPPLEVFCVFSSDFQVPE 169	
RESULT	12		
ID	Q9Z0C4	PRELIMINARY; PRT; 208 AA.	
AC	Q9Z0C4;		
DT	09JUN2004	(TrEMBLrel. 10. Created)	
DT	01-MAY-1999	(TrEMBLrel. 10. Last sequence update)	
DT	01-MAY-2000	(TrEMBLrel. 13. Last annotation update)	

DE C1 AND C4 GENES, CLONE ABURA3-1, PARTIAL  
 GN C1 tobacco leaf curl virus  
 OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=67762;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=ABURA3;  
 RA Ooi K., Ohshita S., Ishii I., Yahara T.;

RT Molecular phylogeny of geminivirus infecting wild plants in Japan.\*;  
 RL J. Plant Res. 110:247-257(1997).  
 RM EMBL: AB001294; BAA33998.1;  
 DR INTERPRO: IPR001191;  
 DR INTERPRO: IPR001301;  
 DR PFAM: PF00799; Gemini\_A1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATLVL1.  
 FT NON\_TER 208 208  
 FT SEQUENCE 208 AA: 23526 MW; 249C31D8729C72D CRC64;

Query Match 71.4%; Score 260; DB 12; Length 208;  
 Best Local Similarity 60.0%; Pred. No. 1.1e-20;  
 Matches 51; Conservative 8; Mismatches 10; Indels 16; Gaps 1;  
 Qy 1 TLVWGFQVDSARGCCOTSNDLLLEALNASSKEALOIIRKPKDFIYFHNLNSL 60  
 Db 102 TLENGTFQVDSARGCCQNDACAEALNASSKEALSIIRKPKDFIYFHNLNSL 161  
 Qy 61 DRI-----FDKTP 69  
 Db 162 DRIFAPLEVPVCPNSSFQDVP 186

Query Match 71.4%; Score 260; DB 12; Length 208;  
 Best Local Similarity 60.0%; Pred. No. 1.1e-20;  
 Matches 51; Conservative 8; Mismatches 10; Indels 16; Gaps 1;

Qy 1 TLVWGFQVDSARGCCOTSNDLLLEALNASSKEALOIIRKPKDFIYFHNLNSL 60  
 Db 102 TLENGTFQVDSARGCCQNDACAEALNASSKEALSIIRKPKDFIYFHNLNSL 161  
 Qy 61 DRI-----FDKTP 69  
 Db 162 DRIFAPLEVPVCPNSSFQDVP 186

RESULT 13  
 Q39180 ID Q39180 PRELIMINARY: PRT: 234 AA.  
 AC Q39180;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DE REPLICATION-DEFECTIVE PROTEIN (FRAGMENT).  
 OS potato yellow mosaic virus.  
 OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10827;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=TONATO STRAIN;  
 RA Ooi K., Ohshita S., Ishii I., Yahara T.,  
 RL J. Plant Res. 110:247-257(1997).  
 RM EMBL: AF026553; RA882603.1;  
 DR INTERPRO: IPR001191;  
 DR INTERPRO: IPR001301;  
 DR PFAM: PF00799; Gemini\_A1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATLVL1.  
 FT NON\_TER 208 208  
 FT SEQUENCE 234 AA: 26486 MW; 9ED8F0697105CD19 CRC64;

Query Match 71.4%; Score 260; DB 12; Length 234;  
 Best Local Similarity 67.1%; Pred. No. 1.2e-20;  
 Matches 47; Conservative 13; Mismatches 10; Indels 0;  
 Qy 1 TLVWGFQVDSARGCCOTSNDLLLEALNASSKEALOIIRKPKDFIYFHNLNSL 60  
 Db 110 TLENGTFQVDSARGCCQNDACAEALNASSKEALSIIRKPKDFIYFHNLSSDL 169  
 Qy 61 DRIFDKTP 70  
 Db 170 DRIFDKTP 179

Query Match 71.4%; Score 260; DB 12; Length 234;  
 Best Local Similarity 67.1%; Pred. No. 1.2e-20;  
 Matches 47; Conservative 13; Mismatches 10; Indels 0;  
 Qy 1 TLVWGFQVDSARGCCOTSNDLLLEALNASSKEALOIIRKPKDFIYFHNLNSL 60  
 Db 110 TLENGTFQVDSARGCCQNDACAEALNASSKEALSIIRKPKDFIYFHNLSSDL 169  
 Qy 61 DRIFDKTP 70  
 Db 170 DRIFDKTP 179

Query Match 71.4%; Score 260; DB 12; Length 234;  
 Best Local Similarity 67.1%; Pred. No. 1.2e-20;  
 Matches 47; Conservative 13; Mismatches 10; Indels 0;  
 Qy 1 TLVWGFQVDSARGCCOTSNDLLLEALNASSKEALOIIRKPKDFIYFHNLNSL 60  
 Db 110 TLENGTFQVDSARGCCQNDACAEALNASSKEALSIIRKPKDFIYFHNLSSDL 169  
 Qy 61 DRIFDKTP 70  
 Db 170 DRIFDKTP 179

Query Match 71.4%; Score 260; DB 12; Length 234;  
 Best Local Similarity 67.1%; Pred. No. 1.2e-20;  
 Matches 47; Conservative 13; Mismatches 10; Indels 0;  
 Qy 1 TLVWGFQVDSARGCCOTSNDLLLEALNASSKEALOIIRKPKDFIYFHNLNSL 60  
 Db 110 TLENGTFQVDSARGCCQNDACAEALNASSKEALSIIRKPKDFIYFHNLSSDL 169  
 Qy 61 DRIFDKTP 70  
 Db 170 DRIFDKTP 179

RESULT 14  
 Q920C0 ID Q920C0 PRELIMINARY: PRT: 208 AA.  
 AC Q920C0;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DE C1 AND C4 GENES, CLONE AMG-1(B152).  
 DE PARTIAL AND COMPLETE CDS (FRAGMENT).  
 GN C1 tobacco leaf curl virus  
 OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=67762;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=AMGB152;  
 RA Ooi K., Ohshita S., Ishii I., Yahara T.;

RT Molecular phylogeny of geminivirus infecting wild plants in Japan.\*;  
 RL J. Plant Res. 110:247-257(1997).  
 RM EMBL: AB001296; BAA33998.1;  
 DR INTERPRO: IPR001191;  
 DR INTERPRO: IPR001301;  
 DR PFAM: PF00799; Gemini\_A1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATLVL1.  
 FT NON\_TER 208 208  
 FT SEQUENCE 208 AA: 23486 MW; E301135F799C3DAD CRC64;

Query Match 70.9%; Score 258; DB 12; Length 208;  
 Best Local Similarity 58.8%; Pred. No. 1.8e-20;  
 Matches 50; Conservative 9; Mismatches 10; Indels 16; Gaps 1;  
 Qy 1 TLVWGFQVDSARGCCOTSNDLLLEALNASSKEALOIIRKPKDFIYFHNLNSL 60  
 Db 102 TLENGTFQVDSARGCCQNDACAEALNASSKEALSIIRKPKDFIYFHNLNSL 161  
 Qy 61 DRI-----FDKTP 69  
 Db 162 DRIFAPLEVPVCPNSSFQDVP 186

Query Match 70.9%; Score 258; DB 12; Length 208;  
 Best Local Similarity 58.8%; Pred. No. 1.8e-20;  
 Matches 50; Conservative 9; Mismatches 10; Indels 16; Gaps 1;  
 Qy 1 TLVWGFQVDSARGCCOTSNDLLLEALNASSKEALOIIRKPKDFIYFHNLNSL 60  
 Db 102 TLENGTFQVDSARGCCQNDACAEALNASSKEALSIIRKPKDFIYFHNLNSL 161  
 Qy 61 DRI-----FDKTP 69  
 Db 162 DRIFAPLEVPVCPNSSFQDVP 186

Query Match 70.9%; Score 258; DB 12; Length 208;  
 Best Local Similarity 58.8%; Pred. No. 1.8e-20;  
 Matches 50; Conservative 9; Mismatches 10; Indels 16; Gaps 1;  
 Qy 1 TLVWGFQVDSARGCCOTSNDLLLEALNASSKEALOIIRKPKDFIYFHNLNSL 60  
 Db 102 TLENGTFQVDSARGCCQNDACAEALNASSKEALSIIRKPKDFIYFHNLNSL 161  
 Qy 61 DRI-----FDKTP 69  
 Db 162 DRIFAPLEVPVCPNSSFQDVP 186

Query Match 70.9%; Score 258; DB 12; Length 208;  
 Best Local Similarity 58.8%; Pred. No. 1.8e-20;  
 Matches 50; Conservative 9; Mismatches 10; Indels 16; Gaps 1;  
 Qy 1 TLVWGFQVDSARGCCOTSNDLLLEALNASSKEALOIIRKPKDFIYFHNLNSL 60  
 Db 102 TLENGTFQVDSARGCCQNDACAEALNASSKEALSIIRKPKDFIYFHNLNSL 161  
 Qy 61 DRI-----FDKTP 69  
 Db 162 DRIFAPLEVPVCPNSSFQDVP 186

Query Match 70.9%; Score 258; DB 12; Length 208;  
 Best Local Similarity 58.8%; Pred. No. 1.8e-20;  
 Matches 50; Conservative 9; Mismatches 10; Indels 16; Gaps 1;  
 Qy 1 TLVWGFQVDSARGCCOTSNDLLLEALNASSKEALOIIRKPKDFIYFHNLNSL 60  
 Db 102 TLENGTFQVDSARGCCQNDACAEALNASSKEALSIIRKPKDFIYFHNLNSL 161  
 Qy 61 DRI-----FDKTP 69  
 Db 162 DRIFAPLEVPVCPNSSFQDVP 186

Query Match 70.9%; Score 258; DB 12; Length 208;  
 Best Local Similarity 58.8%; Pred. No. 1.8e-20;  
 Matches 50; Conservative 9; Mismatches 10; Indels 16; Gaps 1;  
 Qy 1 TLVWGFQVDSARGCCOTSNDLLLEALNASSKEALOIIRKPKDFIYFHNLNSL 60  
 Db 102 TLENGTFQVDSARGCCQNDACAEALNASSKEALSIIRKPKDFIYFHNLNSL 161  
 Qy 61 DRI-----FDKTP 69  
 Db 162 DRIFAPLEVPVCPNSSFQDVP 186

Query Match 70.9%; Score 258; DB 12; Length 208;  
 Best Local Similarity 58.8%; Pred. No. 1.8e-20;  
 Matches 50; Conservative 9; Mismatches 10; Indels 16; Gaps 1;  
 Qy 1 TLVWGFQVDSARGCCOTSNDLLLEALNASSKEALOIIRKPKDFIYFHNLNSL 60  
 Db 102 TLENGTFQVDSARGCCQNDACAEALNASSKEALSIIRKPKDFIYFHNLNSL 161  
 Qy 61 DRI-----FDKTP 69  
 Db 162 DRIFAPLEVPVCPNSSFQDVP 186

Query Match 70.9%; Score 258; DB 12; Length 208;  
 Best Local Similarity 58.8%; Pred. No. 1.8e-20;  
 Matches 50; Conservative 9; Mismatches 10; Indels 16; Gaps 1;  
 Qy 1 TLVWGFQVDSARGCCOTSNDLLLEALNASSKEALOIIRKPKDFIYFHNLNSL 60  
 Db 102 TLENGTFQVDSARGCCQNDACAEALNASSKEALSIIRKPKDFIYFHNLNSL 161  
 Qy 61 DRI-----FDKTP 69  
 Db 162 DRIFAPLEVPVCPNSSFQDVP 186

Query Match 70.9%; Score 258; DB 12; Length 208;  
Best Local Similarity 58.9%; Pred. No. 1.8e-20;  
Matches 50; Conservative 9; Mismatches 10; Indels 16; Gaps 1;  
QY 1 TLVGEFQVDGRSARGGQTSNDLLLEALNASSKEALQIIREKIPKYLFOFHNLSNL 60  
DB 102 TLEWGTQIDGRSARGGQNDACAEALNASSKADALAIIREKLPKDFIOYHNLSNL 161  
QY 61 DRI-----FDKTPF 69  
DB 162 DRIFAPLEVFVCFPSSSFDQVPE 186

Search completed: February 3, 2001, 02:22:53  
Job time: 1890 sec



GenCore version 4.5

Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:24:29 : Search time 81.07 seconds  
(without alignments)  
27.213 Million cell updates/sec

Title: US-09-289-346a-10

Perfect score: 364

Sequence: 1 TLVWGFQVDSRGSGCOT.....FQFHNSLNDRIFDKTPEP 70

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39;+

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	349	95.9	352	1 VAL1_TGMV	P03567 tomato gold
2	249	68.4	361	1 VAL1_PYMV	P27258 potato yell
3	242	66.5	358	1 VAL1_CVVK	P14982 cassava lat
4	242	66.5	358	1 VAL1_CVVK	P14982 cassava lat
5	235	64.6	362	1 VAL1_TVICA	P36275 castor oil
6	228	62.6	349	1 VAL1_TVICA	P36275 castor oil
7	228	62.6	359	1 VAL1_TVLCU	P06523 pepper huas
8	221	60.7	353	1 VAL1_TGMV	P38609 tomato yell
9	221	60.7	359	1 VAL1_TGMV	P05175 bean golden
10	220	60.4	355	1 VAL1_ABYW	P21947 abutilon mo
11	217	59.6	368	1 VAL1_BGVV	P14991 beet curly
12	217	59.6	368	1 VAL1_BGVV	P14991 beet curly
13	206	56.6	357	1 VAL1_TVLCU	P27258 castor oil
14	125	34.3	347	1 VAL1_SLCV	P29048 squash leaf
15	63	17.3	342	1 RSMQ_ECOLI	P39406 escherichia
16	62	17.0	295	1 VAL1_TYDV	P16117 tobacco yell
17	62	17.0	387	1 Y4PF_ARISN	P55615 thizobium s
18	61	16.8	355	1 GLPD_BACSU	P18158 bacillus su
19	61	16.8	355	1 GLPD_BACSU	P18158 bacillus su
20	60	16.5	470	1 RUSK_RHME	P09092 homo sapien
21	60	16.5	799	1 AF5K_STRGO	P54741 streptomyce
22	60	16.5	807	1 AF5K_STRGO	P54742 streptomyce
23	60	16.3	1117	1 TERT_TERTH	P07448 tetrahymena
24	59.5	16.3	79	1 AKHD_DROME	P17975 drosophila
25	58.5	16.1	323	1 VAL1_PASVK	Q00338 panicle str
26	58.5	16.1	1297	1 METL_RAT	P13444 rattus norv
27	58.5	16.1	1297	1 METL_RAT	P13444 rattus norv
28	58.5	16.1	2161	1 Q01658	Q01658 mesocricetu
29	58.5	16.1	2203	1 CCAD_RAT	P27732 rattus norv
30	58.5	15.9	256	1 YKJ9_YEAST	P34247 saccharomyce
31	58	15.9	354	1 CARA_METJA	Q58425 methanococc
32	58	15.9	543	1 NFL_HUMAN	P07196 homo sapien
33	58	15.9	1287	1 SK12_YEAST	P35207 saccharomyce

34 57 15.7 784 1 LON\_RICPR  
35 57 15.7 970 1 DA81\_YEAST  
36 57 15.7 1517 1 RPOC\_CAMEJ  
37 56.5 15.5 266 1 EFEB\_AMYLE  
38 56.5 15.5 295 1 EFEB\_AMYLE  
39 56.5 15.5 295 1 EFEB\_AMYLE  
40 56 15.4 247 1 YCP4\_YEAST  
41 56 15.4 299 1 Y175\_HELPJ  
42 56 15.4 336 1 Y625\_METJA  
43 56 15.4 618 1 MPA\_RAT  
44 56 15.4 624 1 ORC2\_DROME  
45 56 15.4 798 1 VP16\_YEAST

## ALIGNMENTS

RESULT 1  
VAL1\_TGMV STANDARD: PRT: 352 AA.  
ID VAL1\_TGMV  
AC VAL1\_TGMV  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE AL1 PROTEIN.  
GN AC1.  
OS Tomato golden mosaic virus (TGMV).  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OC  
RA Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.:  
"Complete nucleotide sequence of the infectious cloned DNA components  
of tomato golden mosaic virus: potential coding regions and regulatory  
sequences.";  
RL EMBL J. 3:2197-2205(1984).  
CC  
CC -- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@sib-sib.ch).  
CC  
CC EMBL: K02029; -: NOT\_ANNOTATED\_CDS.  
DR PIR: A04170; OOCVLL.  
DR INTERPRO: IPR001191; -.  
DR INTERPRO: IPR001301; -.  
DR PFAM: PF00759; Gemin1\_Alf1.  
DR PRINTS: PR00227; GEMINOTALL1.  
DR PROSITE: PS00226; GEMINOTALLV1.  
KW ATP-binding. 223 230 ATP (POTENTIAL)  
FT NP-BIND 223 230  
SQ SEQUENCE 352 AA: C33C938B5644B4A CRC64:  
Query Match 95.9% Score 349; DB 1; Length 352;  
Best Local Similarity 95.7%; Pred. No. 2.le-31;  
Matches 67; Conservative 5; Indels 0; Gaps 0;  
Oy 1 TLVWGFQVDSRGSGCOTSNLLLEAANASSKEALQIIREKIPKYLQFPHLNSNL 60  
Db 111 TLVWGFQVDSRGSGCOTSNLLLEAANASSKEALQIIREKIPKYLQFPHLNSNL 170  
Oy 61 DRIFDKTPEP 70  
Db 171 DRIFDKTPEP 180  
RESULT 2  
VAL1\_PYMV STANDARD: PRT: 361 AA.

AC P27258;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DE ALL PROTEIN (Rel. 23, Last annotation update)  
 OS Potato yellow mosaic virus (isolate Venezuela).  
 OS Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE:91311403; PubMed:1856690;  
 RA Couteis R.H.A.; Corfin R.S.; Roberts E.J.F.; Hamilton W.D.O.;  
 RA Nucleotide sequence of the infectious cloned DNA components of  
 RT potato yellow mosaic virus.  
 RL J. Gen. Virol. 72:1515-1520(1991).  
 CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: D00940; BAA00782.1;  
 DR PIR: J00364; Q0CVPT.  
 DR INTERPRO: IPR001301;  
 DR PPAM: PF00759; GeminI.ALL: 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR PRINTS: PR00228; GEMCOATCLVL1.  
 KW ATP-binding. 223  
 KW NP-BIND 361 AA; 40850 MW; 5627A33BF1264383 CRC64;  
 FT SEQUENCE 361 AA; 40850 MW; 5627A33BF1264383 CRC64;  
 SQ

Query Match 58.4%; Score 249; DB 1; Length 361;  
 Best Local Similarity 66.7%; Pred. No. 2.4e-20;  
 Matches 46; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

OY 1 TLVGEFQVDRSGRGQTSNDLLLEALNASSKEALQIREKIPKYLFOFHNSNL 60  
 DB 110 TLMGQFQVDRSGRGQTSNDLLLEALNASSKEALQIREKIPKYLFOFHNSNL 169  
 OY 61 DRIFDKTPE 69  
 DB 170 DRIFDKAPE 178

RESULT 3  
 ID VAL1.CLVK STANDARD; PRT: 358 AA.  
 AC P14982;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DE 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE ALL PROTEIN (40.4 KDA PROTEIN).  
 OS Cassava latent virus (strain West Kenyan 844).  
 OS Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RA Stanley J.; Gay M.R.;  
 RA Nucleotide sequence of cassava latent virus DNA.\*;  
 RT Nature 301:260-262(1983).  
 CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: J02057; -; NOT\_ANNOTATED\_CDS.  
 DR INTERPRO: IPR001301;  
 DR PPAM: PF00759; GeminI.ALL: 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR PRINTS: PR00228; GEMCOATCLVL1.  
 KW ATP-binding. 220  
 KW NP-BIND 358 AA; 40346 MW; ED173E753E92D69 CRC64;  
 FT SEQUENCE 358 AA; 40346 MW; ED173E753E92D69 CRC64;  
 SQ

Query Match 66.5%; Score 242; DB 1; Length 358;  
 Best Local Similarity 61.4%; Pred. No. 1.4e-19;  
 Matches 43; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

OY 1 TLVGEFQVDRSGRGQTSNDLLLEALNASSKEALQIREKIPKYLFOFHNSNL 60  
 DB 109 TLMGQFQVDRSGRGQTSNDLLLEALNASSKEALQIREKIPKYLFOFHNSNL 168  
 OY 61 DRIFDKTPE 70  
 DB 169 DRIFQEPAP 178

RESULT 4  
 ID VAL1.CLVK STANDARD; PRT: 358 AA.  
 AC P14972;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DE 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE ALL PROTEIN (40.4 KDA PROTEIN).  
 OS Cassava latent virus (strain Nigerian).  
 OS Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RA Morris B.; Coates L.; Lowe S.; Richardson K.; Eddy P.;  
 RA Nucleotide sequence of the infectious cloned DNA components of  
 RT cassava latent virus (strain Nigerian).  
 RL Nucleic Acids Res 18:197-198(1990).  
 CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: X17095; CA34953.1;  
 DR PIR: S07594; S07594.  
 DR INTERPRO: IPR001191;  
 DR PPAM: PF00759; GeminI.ALL: 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR PRINTS: PR00228; GEMCOATCLVL1.  
 KW ATP-binding. 220  
 KW NP-BIND 358 AA; 40435 MW; 1DB16B0CB2D5E2C CRC64;  
 FT SEQUENCE 358 AA; 40435 MW; 1DB16B0CB2D5E2C CRC64;  
 SQ

Query Match 66.5%; Score 242; DB 1; Length 358;  
 Best Local Similarity 61.4%; Pred. No. 1.4e-19;  
 Matches 43; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

OY 1 TLVGEFQVDRSGRGQTSNDLLLEALNASSKEALQIREKIPKYLFOFHNSNL 60  
 DB 109 TLMGQFQVDRSGRGQTSNDLLLEALNASSKEALQIREKIPKYLFOFHNSNL 168

```

Oy 61 DRIFDTPPEP 70
Db 169 DRIFQEPFAP 178

RESULT 5
VALL_TYICA
ID VALL_TYICA STANDARD: PRT: 362 AA.
AC Q06279;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE ALL PROTEIN (C1 PROTEIN).
GN C1.

OS Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.
RN
RR MEDLINE FROM N.A.
RX PUBMED-8423446;
BA DRY I.B. Sigden, J.P. "Snake I.B.", Mullineaux P.M., Rezaian M.A.;
RT "Nucleotide sequence and genome organization of tomato leaf curl
geminivirus.";
RL J. Gen. Virol. 74:147-151(1993).
CC -1. SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL: S52251; C880888.1; -.
DR EMBL: S52251; C880888.1; -.
DR EMBL: S52251; C880888.1; -.
DR INTERPRO: IPR001191; -.
DR INTERPRO: IPR001191; -.
DR PFAM: PF00759; Gemin_L1; 1.
DR PRINTS: PR00227; GEMCOTALL1.
DR PRINTS: PR00228; GEMCOTCLVL1.
KW ATP-binding. 221 228
KW NP_BIND
KW ATP (POTENTIAL)
FT SEQUENCE 362 AA: 41197 MW: 3432718484704098 CMC64;
Query Match 64.6%; Score 235; DB 1; Length 362;
Best Local Similarity 54.1%; Pred. No. 8.6e-19;
Matches 46; Conservative 9; Mismatches 14; Indels 16; Gaps 1;

Oy 1 TLVNGFQVGRSGCGCTSNLLLELNASSKEEALQIIRKIPKLYLQFPHNLNSLI 60
Db 110 TLENGFQVGRSGCGCTSNLLLELNASSKEEALQIIRKIPKLYLQFPHNLNSLI 169

Oy 61 DRI
Db 170 DRIFTPPEVSPPLSSSFQRPVE 194

RESULT 6
VALL_PHVU
ID VALL_PHVU STANDARD: PRT: 349 AA.
AC Q06923;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE ALL PROTEIN.
GN ALL PROTEIN.
OS Pepper huasteco virus (PHV).
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.
RN
RR MEDLINE FROM N.A.
RX PUBMED-8405944;

Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,
Rivera-Bustamante R.F.;
"Complete nucleotide sequence of pepper huasteco virus: analysis and
comparison with bipartite geminiviruses.";
J. Gen. Virol. 74:2225-2231(1993).
CC -1. SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL: X70418; CAA49855.1; -.
DR EMBL: X70418; CAA49855.1; -.
DR PIR: S31875; S31875.
DR PIR: JQ2300; JQ2300.
DR INTERPRO: IPR001191; -.
DR INTERPRO: IPR001191; -.
DR PFAM: PF00759; Gemin_L1; 1.
DR PRINTS: PR00227; GEMCOTALL1.
DR PRINTS: PR00228; GEMCOTCLVL1.
KW ATP-binding. 221 228
KW NP_BIND
KW ATP (BY SIMILARITY).
FT SEQUENCE 349 AA: 35722 MW: D5F4876CD56370F4 CMC64;
Query Match 62.6%; Score 228; DB 1; Length 349;
Best Local Similarity 60.0%; Pred. No. 4.9e-18;
Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

Oy 1 TLVNGFQVGRSGCGCTSNLLLELNASSKEEALQIIRKIPKLYLQFPHNLNSLI 60
Db 110 TVENGFOVGRSGCGCTSNLLLELNASSKEEALQIIRKIPKLYLQFPHNLNSLI 169

Oy 61 DRIFDTPPEP 70
Db 170 NRIFDTPPEP 179

RESULT 7
VALL_TYICU
ID VALL_TYICU STANDARD: PRT: 359 AA.
AC Q06923;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE ALL PROTEIN (C1 PROTEIN).
GN C1.
OS Tomato yellow leaf curl virus (strain Murcia) (TYLCV).
OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.
RN
RR MEDLINE FROM N.A.
RX MEDLINE-94256836; PubMed-8198442;
BA Norris E., Hidaigo E., Accotto G., Moriones E.;
RT "High similarity among the tomato yellow leaf curl virus isolates
from the west Mediterranean basin: the nucleotide sequence of an
infectious clone from Spain";
Mol. Cell. Probes 8:353-356(1994).
CC -1. SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL: Z25751; CAA41026.1; -.
DR EMBL: Z25751; CAA41026.1; -.
DR PIR: S39211; S39211.
DR INTERPRO: IPR001191; -.

```

```

DR INTERPRO: IPR001301; -
DR PFAM: PF00799; Gemini_All; 1.
DR PRINTS: PR00227; GEMCOATALL.
DR ATP-binding: 221 228 ATP (POTENTIAL).
FT NP_BIND 359 AA: 41065 MW; 2d170a51ef80a3ec CRC64;
SQ SEQUENCE 359 AA: 41065 MW; 2d170a51ef80a3ec CRC64;

Query Match
Best Local Similarity 59.4%; Score 228; DB 1; Length 359;
Matches 41; Conservative 9; Mismatches 19; Indels 0; Gaps 0;

QY 2 LWVGEFQVDSRGAGCGQTNDLLLEALNASKKEALQIREKIPKYLQFPHNLSNLD 61
DB 111 LGNGFTQDGRSGAGGQOTDAYAKAINAGSKSEALDVIKELAPROYLILRHFNINSNLD 170

QY 62 RIFDKTPEP 70
DB 171 KVFQVPAP 179

RESULT 8
ID VALL_BGMV STANDARD: PRT: 353 AA.
AC P05175; 1987 (Rel. 05, Created)
AT 13-AUG-1987 (Rel. 23, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE ALL PROTEIN (40.2 KDA PROTEIN).
GN AC1.
OS Bean golden mosaic virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [J] SEQUENCE FROM N.A.
RA Howarth A.J., Calton J., Bossert M., Goodman R.M.;
RT "Nucleotide sequence of bean golden mosaic virus and a model for gene
regulation in geminiviruses."
RL Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch)
CC EMBL: M10070; AAA46318.1; -
CC INTERPRO: IPR001191; -
CC PFAM: PF00799; Gemini_All; 1.
CC PRINTS: PR00227; GEMCOATALL.
CC ATP-binding: 222 229 ATP (POTENTIAL).
FT NP_BIND 353 AA: 40190 MW; 80fa779df6029a34 CRC64;
SQ SEQUENCE 353 AA: 40190 MW; 80fa779df6029a34 CRC64;

Query Match
Best Local Similarity 60.7%; Score 221; DB 1; Length 353;
Matches 42; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

QY 1 LWVGEFQVDSRGAGCGQTNDLLLEALNASKKEALQIREKIPKYLQFPHNLSNLD 60
DB 111 LGNGFTQDGRSGAGGQOTDAYAKAINAGSKSEALDVIKELAPROYLILRHFNINSNLD 169

QY 61 DRIFDKTPEP 70
DB 170 BRIFKVPAP 179

Query Match
Best Local Similarity 60.7%; Score 221; DB 1; Length 359;
Matches 39; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

QY 2 LWVGEFQVDSRGAGCGQTNDLLLEALNASKKEALQIREKIPKYLQFPHNLSNLD 61
DB 111 LGNGFTQDGRSGAGGQOTDAYAKAINAGSKSEALDVIKELAPROYLILRHFNINSNLD 170

QY 62 RIFDKTPEP 70
DB 171 KVFQVPAP 179

RESULT 10
ID VALL_ABMWV STANDARD: PRT: 355 AA.
AC P1518; 1991 (Rel. 19, Created)
AT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE ALL PROTEIN.
GN AC1.
OS Abutilon mosaic virus (isolate West India).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [J] SEQUENCE FROM N.A.
RA Frischmuth T., Zilmuth G., Jeske H.;
RT "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic
as well as eukaryotic features."
RL Virology 178:461-469(1990).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use, but the statement is long as possible. Content is for commercial
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X15983; -; NOT_ANNOTATED_CDS.
DR EMBL: X04144; -; NOT_ANNOTATED_CDS.
DR INTERPRO: IPR001191; -.
DR INTERPRO: IPR001301; -.
DR PFAM: PF00759; Gemin1_A1.1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00228; GEMCOATCLV1.
KW ATP-binding.
KW NP-BIND.
FT NE-HIND 221 228 ATP (POTENTIAL).
FT SEQUENCE 335 AA: 40257 MW: 1642CARG3251E95 CRC64:
Query Match 60.4%; Score 220; DB 1; Length 355;
Best Local Similarity 58.6%; Pred. No. 3.8e-17;
Matches 41; Conservative 11; Mismatches 18; Indels 0; Gaps 0;
QY 1 TLVNGEFOVDSRGSGCOTSDNLLLEALNASSKEALQIREKIPKYLQFPHLNSML 60
DB 110 TLVNGEFOVDSRGSGCOTSDNLLLEALNASSKEALQIREKIPKYLQFPHLNSML 169
QY 61 DRIFDKTPEP 70
DB 170 ERIFAKAPEP 179
RESULT 11
VAL1_BCTV
ID VAL1_BCTV STANDARD: PRT; 358 AA.
AC P14951;
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE 110 TLVNGEFOVDSRGSGCOTSDNLLLEALNASSKEALQIREKIPKYLQFPHLNSML 169
OC This SWISS-PROT entry is copyright. It is produced through a collaboration
OC between the Swiss Institute of Bioinformatics and the EMBL outstation -
OC the European Bioinformatics Institute. There are no restrictions on its
OC use by non-profit institutions as long as its content is in no way
OC modified and this statement is not removed. Usage by and for commercial
OC entities requires a license agreement (See http://www.isb-sib.ch/announce/
OC or send an email to license@isb-sib.ch).
OC -----
DR EMBL: X04144; -; NOT_ANNOTATED_CDS.
DR INTERPRO: IPR001191; -.
DR INTERPRO: IPR001301; -.
DR PFAM: PF00759; Gemin1_A1.1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00228; GEMCOATCLV1.
KW ATP-binding.
KW NP-BIND.
FT NE-HIND 222 229 ATP (POTENTIAL).
FT SEQUENCE 358 AA: 40889 MW: 39M45E3CB9C333 CRC64:
Query Match 59.6%; Score 217; DB 1; Length 359;
Best Local Similarity 55.7%; Pred. No. 8.3e-17;
Matches 41; Conservative 11; Mismatches 18; Indels 0; Gaps 0;
QY 1 TLVNGEFOVDSRGSGCOTSDNLLLEALNASSKEALQIREKIPKYLQFPHLNSML 60
DB 110 TLVNGEFOVDSRGSGCOTSDNLLLEALNASSKEALQIREKIPKYLQFPHLNSML 169
QY 61 DRIFDKTPEP 70
DB 170 ERIFAKAPEP 179
RESULT 12
VAL1_TM0V
ID VAL1_TM0V STANDARD: PRT; 361 AA.
AC Q06657;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE ALL PROTEIN.
OC This SWISS-PROT entry is copyright. It is produced through a collaboration
OC between the Swiss Institute of Bioinformatics and the EMBL outstation -
OC the European Bioinformatics Institute. There are no restrictions on its
OC use by non-profit institutions as long as its content is in no way
OC modified and this statement is not removed. Usage by and for commercial
OC entities requires a license agreement (See http://www.isb-sib.ch/announce/
OC or send an email to license@isb-sib.ch).
OC -----
DR EMBL: L14460; AAC32414.1; -.
DR PIR: J01870; JQ1870.
DR INTERPRO: IPR001191; -.
DR INTERPRO: IPR001301; -.
DR PFAM: PF00759; Gemin1_A1.1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00228; GEMCOATCLV1.
KW ATP-binding.
KW NP-BIND.
FT NE-BIND 222 229 ATP (BY SIMILARITY).
FT SEQUENCE 361 AA: 40516 MW: 8138B65CE9C6950 CRC64:
Query Match 59.6%; Score 217; DB 1; Length 361;
Best Local Similarity 55.7%; Pred. No. 8.4e-17;
Matches 39; Conservative 15; Mismatches 16; Indels 0; Gaps 0;
QY 1 TLVNGEFOVDSRGSGCOTSDNLLLEALNASSKEALQIREKIPKYLQFPHLNSML 60
DB 110 TLVNGEFOVDSRGSGCOTSDNLLLEALNASSKEALQIREKIPKYLQFPHLNSML 169
QY 61 DRIFDKTPEP 70
DB 170 ERIFAKAPEP 179
RESULT 13
VAL1_TYLCV
ID VAL1_TYLCV STANDARD: PRT; 357 AA.
AC P37259;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE ALL PROTEIN (CI PROTEIN).
DE ALL PROTEIN (CI PROTEIN).
CN CI.

```

```

Matches 39; Conservative 15; Mismatches 16; Indels 0; Gaps 0;
QY 1 TLVNGEFOVDSRGSGCOTSDNLLLEALNASSKEALQIREKIPKYLQFPHLNSML 60
DB 110 TLVNGEFOVDSRGSGCOTSDNLLLEALNASSKEALQIREKIPKYLQFPHLNSML 169
QY 61 DRIFDKTPEP 70
DB 170 ERIFAKAPEP 179
RESULT 12
VAL1_TM0V
ID VAL1_TM0V STANDARD: PRT; 361 AA.
AC Q06657;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE ALL PROTEIN.
OC This SWISS-PROT entry is copyright. It is produced through a collaboration
OC between the Swiss Institute of Bioinformatics and the EMBL outstation -
OC the European Bioinformatics Institute. There are no restrictions on its
OC use by non-profit institutions as long as its content is in no way
OC modified and this statement is not removed. Usage by and for commercial
OC entities requires a license agreement (See http://www.isb-sib.ch/announce/
OC or send an email to license@isb-sib.ch).
OC -----
DR EMBL: L14460; AAC32414.1; -.
DR PIR: J01870; JQ1870.
DR INTERPRO: IPR001191; -.
DR INTERPRO: IPR001301; -.
DR PFAM: PF00759; Gemin1_A1.1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00228; GEMCOATCLV1.
KW ATP-binding.
KW NP-BIND.
FT NE-BIND 222 229 ATP (BY SIMILARITY).
FT SEQUENCE 361 AA: 40516 MW: 8138B65CE9C6950 CRC64:
Query Match 59.6%; Score 217; DB 1; Length 361;
Best Local Similarity 55.7%; Pred. No. 8.4e-17;
Matches 39; Conservative 15; Mismatches 16; Indels 0; Gaps 0;
QY 1 TLVNGEFOVDSRGSGCOTSDNLLLEALNASSKEALQIREKIPKYLQFPHLNSML 60
DB 110 TLVNGEFOVDSRGSGCOTSDNLLLEALNASSKEALQIREKIPKYLQFPHLNSML 169
QY 61 DRIFDKTPEP 70
DB 170 ERIFAKAPEP 179
RESULT 13
VAL1_TYLCV
ID VAL1_TYLCV STANDARD: PRT; 357 AA.
AC P37259;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE ALL PROTEIN (CI PROTEIN).
DE ALL PROTEIN (CI PROTEIN).
CN CI.

```

```

OS Tomato yellow leaf curl virus (TYLCV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN
RP SEQUENCE FROM N.A.
RA MEDLINE-9102449; PubMed-1926771;
RA Navot N, Pichersky E, Zaidan M, Zamir D, Czosnek H;
RT "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
RT with a single genomic component.";
RL Virology 185:151-161(1991).
CC
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X15656; CA33688.1; -.
DR MEDLINE-9102449; PubMed-1926771;
DR Navot N, Pichersky E, Zaidan M, Zamir D, Czosnek H;
DR INTERPRO: IPRO01191; -.
DR PFAM: PF00799; Gemin_L1;
DR PRINTS; PR00227; GEMCOATL1.
DR ATPT-binding; 218 225 ATP (POTENTIAL).
KW NP-BIND
SQ SEQUENCE 357 AA; 40678 MW; 939A668E1A382A7 CRC64;

Query Match 56.6%; Score 206; DB 1; Length 357;
Best Local Similarity 63.9%; Pred. No. 1.4e-15;
Matches 39; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 4 WGFQVDSRGAGCGCTNDLLLEALNASKKEALQITREKIPKYLPOPHNLSNLDRI 63
DB 111 FVSQIDGSRAGCGQSSANDAAEALNLSKSEALNLTAKRPDYILQPHNLSNLDRI 170

QY 64 F 64
DB 171 F 171

RESULT 14
VAL1_SUCV STANDARD; PRT; 347 AA.
ID VAL1_SUCV
IC P29048;
DT 01-DEC-1992 (Rel. 24; Created)
DT 01-DEC-1992 (Rel. 24; Last sequence update)
DT 01-DEC-1992 (Rel. 24; Last annotation update)
DR ALL PROTEIN
OS Squash leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN
RP SEQUENCE FROM N.A.
RA MEDLINE-9102449; PubMed-1984568;
RA Lazarowitz S.G., Lazdins J.B.;
RA "Components of a bipartite squash leaf curl geminivirus with a broad
RT host range phenotype.";
RL Virology 180:58-69(1991).
CC
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M38183; AAC32410.1; ALT-INIT.

PIR; C36785; QOCVSL.
DR INTERPRO: IPRO01191; -.
DR PFAM; PF00799; Gemin_L1;
DR PRINTS; PR00227; GEMCOATL1.
KW ATP-binding 218 225 ATP (POTENTIAL).
FT NP-BIND
SQ SEQUENCE 347 AA; 39110 MW; AFDABDE122110E CRC64;

Query Match 34.3%; Score 125; DB 1; Length 347;
Best Local Similarity 37.9%; Pred. No. 1.2e-06;
Matches 25; Conservative 15; Mismatches 22; Indels 4; Gaps 1;

QY 5 GEFQVDSRGAGCGCTNDLLLEALNASKKEALQITREKIPKYLPOPHNLSNLDRI 64
DB 116 GQYKVSQ---GSKSKDDYINAVNAGSAGALDLKAGDKPTIVNYINILANVERLF 171

QY 65 DKTPEP 70
DB 172 QKPPRP 177

RESULT 15
RSMC_ECOLI STANDARD; PRT; 342 AA.
ID RSMC_ECOLI
IC MEDLINE-95334362; PubMed-7610040;
DT 01-FEB-1995 (Rel. 31; Created)
DT 01-FEB-1995 (Rel. 31; Last sequence update)
DT 01-OCT-2000 (Rel. 40; Last annotation update)
DE RIBOSOMAL RNA SMALL SUBUNIT METHYLTRANSFERASE C (EC 2.1.1.52) (RNA
DE [GUANINE-N2]-METHYLTRANSFERASE) (16S RNA M2G1207 METHYLTRANSFERASE).
GN RSMC.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN
RP SEQUENCE FROM N.A.
RA MEDLINE-95334362; PubMed-7610040;
RA Hurland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
RA Plattner F.R.;
RA "Analysis of the Escherichia coli genome VI: DNA sequence of the
RT region from 92.8 through 100 minutes.";
RL Nucleic Acids Res. 23:2105-2119(1995).
RN
RP SEQUENCE OF 1-19, AND CHARACTERIZATION.
RA Tscherne J.S., Nurse K., Popenick P., Ofengand J.;
RA "Purification, cloning, and characterization of the 16S RNA M2G1207
RA methyltransferase from Escherichia coli.";
RA Nucleic Acids Res. 27:5925-5932(1999).
CC
CC -!- FUNCTION: SPECIFICALLY METHYLATES THE GUANOSINE IN POSITION 1207
CC OF 16S RNA IN THE 30S PARTICLE.
CC -!- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + RNA = S-ADENOSYL-L-
CC HOMOCYSTEINE + RNA CONTAINING N2-METHYLGUANINE.
CC -!- COFACTOR: REQUIRES MAGNESIUM FOR ACTIVITY.
CC -!- SIMILARITY: BELONGS TO THE METHYLTRANSFERASE SUPERFAMILY. RSMC
CC SUPERFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U14003; AAA97267.1; -.
DR EMBL; AE000507; AAC77324.1; -.
DR ECGENE; EG12596; RSMC.
DR PROSITE; PS00092; NC_M7PASE; UNKNOWN_1.
KW RNA processing; Transferase; Methyltransferase; Magnesium.

```

FT INIT\_MET 0  
SQ SEQUENCE 342 AA: 37493 MW: C7A318155700302D CRG64;

Query Match 17 3%; Score 63; DB 1; Length 342;  
Best Local Similarity 31.1%; Pred. No. 8.2;  
Matches 14; Conservative 7; Mismatches 16; Indels 8; Gaps 1;

QY 4 WGEKQVDGRS-----ARGGQTSNDLLLEALNASSKEALQI 40  
DB 158 WGEISVDGLVKTLEGVFSRGLDVGSQLLSLTTPHKGKLDV 202

Search completed: February 3, 2001, 02:24:30  
Job time: 642 sec

**THIS PAGE BLANK (USPTO)**



GenCorte version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:17:41 ; Search time 118.74 Seconds  
(without alignments)  
40.029 Million cell updates/sec

Title: US-09-289-346a-10

Perfect score: 364

Sequence: 1 TLVWGFQVDSRGSGCQT.....FOFHNSNLDRIFDKTEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900635 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR66:\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	349	95.9	352	1 QOCV11	All protein - toma
2	249	68.4	361	1 QOCVPT	All protein - toma
3	249	68.4	361	1 QOCVPT	hypothetical prote
4	235	66.5	358	2 S07594	All protein - toma
5	228	62.6	349	2 J02300	replicase - pepper
6	228	62.6	349	2 S11875	All protein - pep
7	228	62.6	349	2 S11875	gene C1 protein - toma
8	221	60.7	359	2 S22595	All protein - toma
9	220	60.4	355	1 QOCVW1	All protein - abu
10	218	59.9	351	2 J02327	All protein - indi
11	217	59.6	358	1 J01870	All protein - toma
12	217	59.6	385	2 S28360	All protein - beet
13	215	59.1	359	2 S29235	gene C1 protein -
14	208	57.1	360	2 S59885	replication-associ
15	208	57.1	360	2 S59885	All protein - toma
16	138	37.9	311	2 S45059	All protein - toma
17	125	34.3	347	1 QOCV81	All protein - squa
18	64	17.6	587	2 J01419	PC gamma (Igc) rec
19	63.5	17.4	1239	2 T48959	kinesin-like prote
20	63	17.3	343	2 S56595	hypothetical 37.6k
21	62	17.0	295	2 D42452	CI protein - tobac
22	61.2	16.9	416	2 A26972	hypothetical prote
23	61.2	16.9	416	2 A26972	hypothetical prote
24	61.5	16.9	481	2 A70091	type IIS restricti
25	61.5	16.9	1279	2 E64709	glycerol-3-phospha
26	61	16.8	555	2 C45868	granulocyte colony
27	61	16.8	771	2 B38252	granulocyte colony
28	61	16.8	783	2 JH0329	granulocyte colony
29	61	16.8	863	2 C38252	granulocyte colony

30 60.5 16.6 447 2 T12544  
31 60.5 16.6 1792 2 T13939  
32 60 16.5 470 2 T46814  
33 60 16.5 1799 2 T48889  
34 60 16.5 1255 2 T02331  
35 60 16.5 1255 2 T02331  
36 59.5 16.3 79 1 A38656  
37 59.5 16.3 210 2 T05021  
38 59.5 16.3 388 2 C69196  
39 59.5 16.3 546 2 A59484  
40 59 16.2 373 2 A16950  
41 59 16.2 506 2 T19287  
42 59 16.2 1766 2 T02331  
43 59 16.2 1766 2 T02331  
44 58.5 16.1 323 2 J01552  
45 58.5 16.1 397 2 S06114

#### ALIGNMENTS

RESULT 1  
QOCV11  
C:Species: tomato golden mosaic virus  
A:Note: host Nicotiana sp. (tobacco)  
C:Date: 28-Aug-1985 sequence\_revision 20-Aug-1985 text\_change 08-Apr-1994  
C:Accession: A01170 Stein, V.E.; Coutts, R.H.A.; Buck, K.W.  
EMBL J3, 2197-2205, 1984  
A:Title: Complete nucleotide sequence of the infectious cloned DNA components of Toma  
A:Reference number: A01153  
A:Accession: A01170  
A:Molecule type: DNA  
A:Residues: 1-352 <HAM>  
C:Comment: The genome consists of two circular, single-stranded DNA components, DNA A  
C:Genetics:  
A:Map position: segment A  
C:Superfamily: tomato golden mosaic virus ALL protein

Query Match 95.9% Score 349 DB 1 Length 352  
Query Similarity 95.7% Identical 31  
Matches 67: Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Oy 1 TLVWGFQVDSRGSGCQTSNLLALNASKEALQIIREKIPEKYLQFHLNSNL 60  
Dbs 111 TLVWGFQVDSRGSGCQTSNLLALNASKEALQIIREKIPEKYLQFHLNSNL 170  
Oy 61 DRIEDKTEP 70  
Dbs 171 DRIEDKTEP 180  
RESULT 2  
QOCVPT  
A:All protein - potato yellow mosaic virus (isolate Venezuela)  
C:Species: potato yellow mosaic virus  
C:Date: 30-Jun-1992 sequence\_revision 30-Jun-1992 text\_change 16-Jun-2000  
C:Accession: J00364  
R:Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.  
J. Gen. Virol. 72, 1515-1520, 1991  
A:Title: The nucleotide sequence of the infectious cloned DNA components of potato ye  
A:Reference number: J00362; MUD:91311403  
A:Accession: J00364  
A:Status: translation not shown  
A:Residues: 1-361 <COQ>  
A:Cross-references: GB:D00940; NID:g222458; PIDN:BAA00782.1; PID:g222459  
C:Genetics:  
A:Map position: segment A  
C:Superfamily: tomato golden mosaic virus ALL protein

```

DB      110  TLEWGFQIDGRSNGQGSNDYIAQAALVTGSKSEALNVLRLAPAKYIVLPQHNLNSWL 169
OY      61  DRI-----EDKTE 69
DB      170  DRIETFPPLVYVSPFLSSSDRPE 194
RESULT  5
J02300  replicase - pepper huasteco virus (component A)
Accession: J02300
Submitter: National Institute of Health
C:Date: 14-Jul-1994
C:Update: 14-Jul-1994
C:Accession: J02300
R:Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante, J. Gen. Virol. 74, 2225-2231, 1993
A:Title: Complete nucleotide of pepper huasteco virus. Analysis and comparison with b
A:Reference number: J02299; MUID:94015007
A:Accession: J02300
A:Molecule type: DNA
A:Size: 8428
A:Accession: J02300
C:Superfamily: tomato golden mosaic virus AII protein
Query Match      62.5%; Score 228; DB 2; Length 349;
Best Local Similarity 60.0%; Pred. No. 6e-18;
Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;
OY      1  TLEWGFQVDGRSNGQGSNDYILLEALNASKKEALQIIRKIPKTYIFQHHNLNSWL 60
DB      110  TLEWGFQIDGRSNGQGSNDYIAQAALVTGSKSEALNVLRLAPAKYIVLPQHNLNSWL 169

```

```

DB      170 NR16TPPPE 179
|-----|-----|-----|-----|-----|-----|
|          |          |          |          |          |          |
| RESULT   6         |          |          |          |          |          |
| S11875   |          |          |          |          |          |
| All protein - pepper rizado amarillo virus |          |          |          |          |          |
| C:Species? Pepper rizado amarillo virus    |          |          |          |          |          |
| C:Species? sequence_revision 26-May-1995 *text_change 20-Sep-1999 |          |          |          |          |          |
| C:Accession: S11875 |          |          |          |          |          |
| R:Torres-Pacheco, I.; Garzon-Tinoco, J.A.; Jimenez, B.; Herrera-Estrella, L.; Rivera |          |          |          |          |          |
| submitted to the EMBL Data Library, February 1993 |          |          |          |          |          |
| A:Description: Complete nucleotide sequence of pepper huasteco virus: analysis and co |          |          |          |          |          |
| A:Reference number: S11872 |          |          |          |          |          |
| A:Accession: S11875 |          |          |          |          |          |
| A:Accession: S11875 |          |          |          |          |          |
| A:Molecule type: DNA |          |          |          |          |          |
| A:Residues: 1-349 <TOR> |          |          |          |          |          |
| A:Cross-references: EMBL:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025 |          |          |          |          |
| A:Note: the source is designated as pepper huasteco virus |          |          |          |          |          |
| C:Superfamily: tomato golden mosaic virus ALL protein |          |          |          |          |          |
|          |          |          |          |          |          |
| Query Match       62.6%; Score 228; DB 2: Length.349; |          |          |          |          |          |
| Best Local Similarity 60.0%; Pred.No. 6e-18; |          |          |          |          |          |
| Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0; |          |          |          |          |          |
|          |          |          |          |          |          |
| QY      1 TLWGEQFVDGRSGRCGTSNDLLLEALNSAKSEALOTIREKIPETLYCFHNHLSML 60 |          |          |          |          |          |
|          |          |          |          |          |          |
| Db      110 TWMECEQTIDGRSARGQGQSANDTYAKALNSASAEALQIKERISQPHFLQFHNVSNR 169

```

Db 170 NRIFOTPEP 179

|||||

RESULT 7

S39211

gene C1 protein - tomato yellow leaf curl virus

C:Species: tomato yellow leaf curl virus  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 20-Sep-1999  
 C:Accession: S39211  
 R:Norris, E.; Hidalgo, E.; Accotto, G.; Moriones, E.  
 A:Description: Tomato yellow leaf curl virus isolate from the  
 A:Description: High similarity among the tomato yellow leaf curl virus isolates from the  
 A:Reference number: S39209  
 A:Accession: S39211  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-359 <NR>  
 A:Cross-references: EMBL:Z25751; NID:g433655; PIDN:CAAB1026.1; PID:g433658  
 C:Superfamily: tomato golden mosaic virus All protein

Query Match 62.6%; Score 228; DB 2: Length 359;  
 Best Local Similarity 59.4%; Pred. No. 6.2e-18;  
 Matches 41; Conservative 9; Mismatches 19; Indels 0; Gaps 0;

QY 2 LWGCFQDGRSARGCQTSNDLLLEALNASSKEALQITREKIPKYLFPFHNSNLD 61  
 Db 111 LWFQTFQDGRSARGCQQTANDAYAKAINAGSKSQALDVIKELAPRVILHFNHNSLD 170

QY 62 RIFDKTPPEP 70  
 Db 171 RYFQVPPAP 179

RESULT 8

hypothetical protein C4 - tomato yellow leaf curl virus  
 C:Species: tomato yellow leaf curl virus  
 C:Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 20-Sep-1999  
 C:Accession: S22593  
 R:Rheyr-Tour, A.; Bendahmane, M.; Matzeit, V.; Accotto, G.P.; Crespi, S.; Gronenborn, B.  
 A:Description: Tomato yellow leaf curl virus isolate from Sardinia is a whitefly-transmitted monopartite  
 A:Title: Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monopartite  
 A:Reference number: S22588; MUID:92107660  
 A:Accession: S22593  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-359 <NR>  
 A:Cross-references: EMBL:X61153; NID:g62211; PIDN:CAAB3466.1; PID:g62217  
 A:Superfamily: tomato golden mosaic virus All protein

Query Match 60.7%; Score 221; DB 2: Length 359;  
 Best Local Similarity 56.5%; Pred. No. 3.8e-17;  
 Matches 39; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

QY 2 LWGCFQDGRSARGCQTSNDLLLEALNASSKEALQITREKIPKYLFPFHNSNLD 61  
 Db 111 LWFQTFQDGRSARGCQQTANDAYAKAINAGSKSQALDVIKELAPRVILHFNHNSLD 170

QY 62 RIFDKTPPEP 70  
 Db 171 RYFQVPPAP 179

RESULT 9

AV1 protein - abutilon mosaic virus (isolate West India)  
 C:Species: abutilon mosaic virus  
 C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 08-Apr-1994  
 C:Accession: A36214  
 R:Frickschmidt, A.; Zilm, G.; Jeske, H.  
 A:Description: Abutilon mosaic virus isolate from West India  
 A:Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as well  
 A:Reference number: A36214; MUID:91020984  
 A:Accession: A36214  
 A:Molecule type: DNA  
 A:Residues: 1-355 <FRI>

A:Cross-references: EMBL:X15983  
 C:Genetics:  
 A:Map position: segment A  
 C:Superfamily: tomato golden mosaic virus All protein

Query Match 60.4%; Score 220; DB 1: Length 355;  
 Best Local Similarity 58.6%; Pred. No. 4.8e-17;  
 Matches 41; Conservative 11; Mismatches 18; Indels 0; Gaps 0;

QY 1 TLVNGFQDGRSARGCQTSNDLLLEALNASSKEALQITREKIPKYLFPFHNSNLD 60  
 Db 110 TAMEGFOIDGRSARGCQQTANDSTAKALNAGDVQSALNLEAEQDQVILQNHRSNLD 169

QY 61 DRIFDKTPPEP 70  
 Db 170 BRIFAKAPEP 179

RESULT 10

QJ2327 protein - Indian cassava mosaic virus  
 N:Alternate names: Replication-associated protein  
 C:Species: Indian cassava mosaic virus  
 C:Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 20-Sep-1999  
 C:Accession: J02327; S35883  
 R:Hong, Y.G.; Robinson, D.J.; Harrisou, B.D.  
 J:Gen. Virol. 74, 2437-2443, 1993  
 A:Title: Nucleotide sequence evidence for the occurrence of three distinct whitefly-t  
 A:Reference number: J02327; MUID:94065670  
 A:Accession: J02327  
 A:Molecule type: DNA  
 A:Residues: 1-351 <NR>  
 A:Cross-references: EMBL:Z24758; NID:g395351; PIDN:CAAB0891.1; PID:g584046  
 C:Superfamily: tomato golden mosaic virus All protein

Query Match 59.9%; Score 218; DB 2: Length 351;  
 Best Local Similarity 61.2%; Pred. No. 7.9e-17;  
 Matches 41; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

QY 4 MGFQTFQDGRSARGCQTSNDLLLEALNASSKEALQITREKIPKYLFPFHNSNLDRI 63  
 Db 113 WFTQTFQDGRSARGCQQTANDAYAKAINAGSKSALNLELAPRYLDFHHSNLDRI 172

QY 64 FQTFKTPPEP 70  
 Db 173 FTKFPPPP 179

RESULT 11

QJ1870 protein - tomato mottle virus (isolate Florida)  
 C:Species: tomato mottle virus  
 C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 07-May-1999  
 C:Accession: J01870  
 R:Abouvid, A.M.; Polston, J.E.; Hiebert, E.  
 J:Gen. Virol. 73, 3225-3229, 1992  
 A:Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated f  
 A:Reference number: J01869; MUID:93107858  
 A:Accession: J01870  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-358 <ABO>  
 A:Cross-references: GB:L14460  
 C:Genetics:

A:Map position: segment A  
 C:Superfamily: tomato golden mosaic virus All protein

Query Match 59.6%; Score 217; DB 1: Length 358;  
 Best Local Similarity 55.7%; Pred. No. 1.1e-16;  
 Matches 39; Conservative 15; Mismatches 16; Indels 0; Gaps 0;

```

QY 1 TLWGEFQVDRSGRGCGQTSDNLLLEALNASKKEALQITREKIPKYLFOFHNLSNL 60
Db 107 TIEWGFIQDRSGRGCGQTSDNLLLEALNASKKEALQITREKIPKYLFOFHNLSNL 166
QY 61 DRIFDKTPEP 70
Db 167 BRIFAKAPEP 176

RESULT 12
Best Local Similarity 59.1%; Score 215; DB 2; Length 359;
Matches 38; Conservative 12; Mismatches 19; Indels 0; Gaps 0;
A1: protein - beet curly top virus
C:Species: beet curly top virus
C>Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 20-Sep-1999
R:Accession: S28360
R:Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.
R:Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly top virus
A:Reference: S28360
A:Accession: S28360
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-385 <STA>
A:Cross-references: EMBL:X04144; NID:g210576; PIDN:AAA42751.1; PID:g210579
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 59.1%; Score 217; DB 2; Length 385;
Best Local Similarity 55.7%; Pred. No. 1.1e-16;
Matches 39; Conservative 15; Mismatches 16; Indels 0; Gaps 0;
QY 1 TLWGEFQVDRSGRGCGQTSDNLLLEALNASKKEALQITREKIPKYLFOFHNLSNL 60
Db 137 TIEWGFIQDRSGRGCGQTSDNLLLEALNASKKEALQITREKIPKYLFOFHNLSNL 196
QY 61 DRIFDKTPEP 70
Db 197 QKIFORPPPP 206

RESULT 13
Best Local Similarity 59.1%; Score 215; DB 2; Length 359;
Matches 38; Conservative 12; Mismatches 19; Indels 0; Gaps 0;
A1: protein - tomato yellow leaf curl virus
C:Species: tomato yellow leaf curl virus
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Sep-1997
R:Accession: S39235
R:Crespi, S.; Norris, E.; Vair, A.; Bosco, D.; Accotto, G.
R:Title: The nucleotide sequence of a full-length infectious clone of the geminivirus tomato yellow leaf curl virus
A:Reference: S39235
A:Accession: S39235
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-359 <CRE>
A:Cross-references: EMBL:g1041671; NID:g1334964
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 59.1%; Score 215; DB 2; Length 359;
Best Local Similarity 55.1%; Pred. No. 1.8e-16;
Matches 38; Conservative 12; Mismatches 19; Indels 0; Gaps 0;
QY 2 LWGEFQVDRSGRGCGQTSDNLLLEALNASKKEALQITREKIPKYLFOFHNLSNLD 61
Db 111 LEMGTFQDRSGRGCGQTSDNLLLEALNASKKEALQITREKIPKYLFOFHNLSNLD 170
QY 62 RIFDKTPEP 70
Db 171 KVFQVPPAP 179

*-RESULT 14

```

```

S59885
replication-associated protein Cl - tomato yellow leaf curl virus
C:Species: tomato yellow leaf curl virus
C>Date: 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 20-Sep-1999
R:Accession: S59885
R:Hoon, Y.; Harrison, B.D.
R:Title: The nucleotide sequence of a full-length infectious clone of the geminivirus tomato yellow leaf curl virus
A:Reference: S59885
A:Accession: S59885
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-360 <HON>
A:Cross-references: EMBL:Z48182; NID:g944838; PIDN:CAA80229.1; PID:g974211
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 57.1%; Score 208; DB 2; Length 360;
Best Local Similarity 53.1%; Pred. No. 1.1e-15;
Matches 39; Conservative 12; Mismatches 15; Indels 0; Gaps 0;
QY 4 WGEFQVDRSGRGCGQTSDNLLLEALNASKKEALQITREKIPKYLFOFHNLSNLD 63
Db 113 FGVIQDRSGRGCGQTSDNLLLEALNASKKEALQITREKIPKYLFOFHNLSNLD 172
QY 64 FDKTPEP 69
Db 173 FTFPSAE 178

RESULT 15
QOCVCI
A1: protein - tomato yellow leaf curl virus
R:Alternate names: Cl protein curl virus
R:Title: The nucleotide sequence of a full-length infectious clone of the geminivirus tomato yellow leaf curl virus
C>Date: 30-Jun-1997 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
R:Accession: D40779
R:Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.
R:Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin
A:Reference: D40779; MUID:92024070
A:Accession: D40779
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-357 <NAV>
A:Cross-references: GB:X15656; NID:g62204; PIDN:CAA33688.1; PID:g62207
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 56.6%; Score 206; DB 1; Length 357;
Best Local Similarity 63.9%; Pred. No. 1.8e-15;
Matches 39; Conservative 10; Mismatches 12; Indels 0; Gaps 0;
QY 4 WGEFQVDRSGRGCGQTSDNLLLEALNASKKEALQITREKIPKYLFOFHNLSNLD 63
Db 111 FGVIQDRSGRGCGQTSDNLLLEALNASKKEALQITREKIPKYLFOFHNLSNLD 170
QY 64 F 64
Db 171 F 171

Search completed: February 3, 2001, 02:17:42
Job time: 2214 sec

```

GenCode version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: February 3, 2001, 02:19:51 ; Search time 109.09 seconds  
(without alignments)  
11.523 Million cell updates/sec

Title: US-09-289-346a-10

Perfect score: 364

Sequence: 1 TLVWGEFQVDSARGSCQT.....POFHNLSNLDRIEDTKPEP 70

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2\_5/ptodata/1/aa/5a/COMB.pep.\*

2: /cgn2\_5/ptodata/1/aa/5b/COMB.pep.\*

3: /cgn2\_5/ptodata/1/aa/6/COMB.pep.\*

4: /cgn2\_5/ptodata/1/aa/PCUTUS.COMB.pep.\*

5: /cgn2\_5/ptodata/1/aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	221	60.7	359	3	US-08-809-103B-2
2	221	60.7	359	3	US-08-809-103B-4
3	221	60.7	359	3	US-08-809-103B-6
4	221	60.7	359	3	US-08-809-103B-8
5	61	16.8	602	2	US-08-419-652-6
6	61	16.8	763	5	5423248-376-6
7	61	16.8	836	1	US-07-923-976-6
8	61	16.8	836	1	US-07-923-976-8
9	61	16.8	863	1	US-07-745-206A-2
10	58.5	16.1	2161	1	US-07-745-206A-4
11	58.5	16.1	2161	1	US-08-455-543A-49
12	58.5	16.1	2161	1	US-08-455-543A-51
13	58.5	16.1	2161	2	US-08-223-305C-49
14	58.5	16.1	2161	2	US-08-223-305C-51
15	58.5	16.1	2161	2	US-08-311-365C-21
16	56	15.4	323	3	US-07-667-276A-7
17	56	15.4	844	3	US-09-029-267-20
18	55	15.1	751	3	US-09-036-987A-24
19	54	14.8	244	1	US-08-910-973-22
20	54	14.8	432	1	US-08-167-919A-11
21	54	14.8	172	3	US-08-113-106-11
22	54	14.8	172	3	US-08-113-106-11
23	54	14.8	1713	4	PCT-US94-102614*24
24	53.5	14.7	439	1	US-07-637-870-9
25	53.5	14.7	439	1	US-07-637-399-6
26	53.5	14.7	439	1	US-08-112-703-6
27	53.5	14.7	3898	2	US-08-876-991-2
28	53.5	14.7	3898	2	US-09-059-853-2

29 53 14.6 126 1 US-08-083-949-1 Sequence 1, Appl  
30 53 14.6 237 1 US-08-910-973-13 Sequence 13, Appl  
31 53 14.6 255 1 US-07-901-707-6 Sequence 6, Appl  
32 53 14.6 255 1 US-07-988-430-6 Sequence 6, Appl  
33 53 14.6 255 1 US-08-444-336-6 Sequence 6, Appl  
34 53 14.6 255 4 PCT-US92-03467-6 Sequence 6, Appl  
35 53 14.6 783 5 5231168-2 Patent No. 5231168  
36 52.5 14.4 770 1 US-08-445-135-2 Sequence 2, Appl  
37 52.5 14.4 870 2 US-08-266-311-1 Sequence 1, Appl  
38 52.5 14.4 870 2 US-08-467-527A-1 Sequence 1, Appl  
39 52.5 14.4 3079 4 PCT-US94-00198-4 Sequence 4, Appl  
40 52 14.3 271 1 US-08-276-919-10 Sequence 10, Appl  
41 52 14.3 271 1 PCT-US95-08513-13 Sequence 13, Appl  
42 52 14.3 271 1 PCT-US95-08513-13 Sequence 13, Appl  
43 52 14.3 274 1 US-08-776-088-15 Sequence 15, Appl  
44 52 14.3 274 4 PCT-US95-09145A-15 Sequence 15, Appl  
45 52 14.3 325 1 US-08-276-919-4 Sequence 4, Appl

#### ALIGNMENTS

RESULT 1  
US-08-809-103B-2  
? Sequence 2, Application US/08809103B  
? GENERAL INFORMATION:  
? APPLICANT: GRONENBORN, Bruno  
? TITLE OF INVENTION: TRANSGENIC DNA VIRUS RESISTANT  
? NUMBER OF SEQUENCES: 17  
? CORRESPONDENCE ADDRESS:  
? ADDRESSER: YOUNG & THOMPSON  
? STREET: 745 South 23rd Street  
? CITY: Arlington  
? STATE: Virginia  
? COUNTRY: U.S.A.  
? ZIP: 22202  
? COMPUTER READABLE FORM:  
? MEDIUM TYPE: Floppy disk  
? COMPUTER: IBM PC Compatible  
? OPERATING SYSTEM: PC-DOS/MS-DOS  
? SOFTWARE: Patentin Release #1.0, Version #1.30  
? COUNTRY OF ORIGIN: USA  
? APPLICATION NUMBER: US/08/809,103B  
? FILING DATE: 17-MAR-1997  
? CLASSIFICATION: 800  
? PRIOR APPLICATION DATA: FR 94.11040  
? APPLICATION NUMBER: 15-SEP-1994  
? FILING DATE: 15-SEP-1994  
? PRIOR APPLICATION DATA: NO PCT/FR95/01192  
? APPLICATION NUMBER: 15-SEP-1995  
? FILING DATE: 15-SEP-1995  
? ATTORNEY/AGENT INFORMATION:  
? NAME: PATCH, Andrew J.  
? REGISTRATION NUMBER: 32,925  
? REFERENCE/DOCKET NUMBER: USB94AL CNR TOM  
? TELECOMMUNICATION INFORMATION:  
? TELEPHONE: (703) 562-21-2297  
? TELEFAX: (703) 562-21-5773  
? TELEX: 248425 EMBON  
? INFORMATION FOR SEQ ID NO: 2:  
? SEQUENCE CHARACTERISTICS:  
? LENGTH: 359 amino acids  
? TYPE: amino acid  
? TOPOLOGY: Linear  
? Molecule Type: protein  
US-08-809-103B-2

Query Match 60.7%; Score 221; DB 3; Length 359;  
Best Local Similarity 56.5%; Pred. No. 2.2e-21;  
Matches 39; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

QY 2 LVNGFQVGRSARGCQTSNDLLLEALNASKEALQIIREKIPKYLQFPHNLSNLD 61  
 Db 111 LVNGFQVGRSARGCQTSNDLLLEALNASKEALQIIREKIPKYLQFPHNLSNLD 170  
 QY 62 RIFDKTPEP 70  
 Db 171 RVFQVPPAP 179

## RESULT 2

US-08-809-103b-4  
 : Sequence 4, Application US/08809103b  
 : Patent No. 6133505  
 : GENERAL INFORMATION:  
 : APPLICANT: GRONENBORN, Bruno  
 : TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT  
 : TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME  
 : NUMBER OF SEQUENCES: 17  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: YOUNG & THOMPSON  
 : STREET: 745 South 23rd Street  
 : CITY: Arlington  
 : STATE: Virginia  
 : COUNTRY: U.S.A.  
 : ZIP: 22202  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: PatentIn Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/809,103b  
 : FILING DATE: 17-MAR-1997  
 : PRIORITY APPLICATION DATA:  
 : PRIOR APPLICATION NUMBER: FR 94.11040  
 : FILING DATE: 15-SEP-1994  
 : APPLICATION TYPE: WO PCT/FR95/01192  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: PATCH, Andrew J.  
 : REGISTRATION NUMBER: 32,925  
 : REFERENCE/DOCKET NUMBER: USB94AL CNR TOM  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: (703) 521-2297  
 : TELEX: 248425 EMBON  
 : FAX: (703) 685-0573  
 : INFORMATION FOR SEQ ID NO: 4:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 359 amino acids  
 : TYPE: amino acid  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: protein  
 : US-08-809-103b-4

Query Match 60.7%; Score 221; DB 3; Length 359;  
 Best Local Similarity 56.5%; Pred. No. 2,2e-21;  
 Matches 39; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

QY 2 LVNGFQVGRSARGCQTSNDLLLEALNASKEALQIIREKIPKYLQFPHNLSNLD 61  
 Db 111 LVNGFQVGRSARGCQTSNDLLLEALNASKEALQIIREKIPKYLQFPHNLSNLD 170  
 QY 62 RIFDKTPEP 70  
 Db 171 RVFQVPPAP 179

## RESULT 3

US-08-809-103b-6

: Sequence 6, Application US/08809103b  
 : Patent No. 6133505  
 : GENERAL INFORMATION:  
 : APPLICANT: GRONENBORN, Bruno  
 : TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT  
 : TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME  
 : NUMBER OF SEQUENCES: 17  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: YOUNG & THOMPSON  
 : STREET: 745 South 23rd Street  
 : CITY: Arlington  
 : STATE: Virginia  
 : COUNTRY: U.S.A.  
 : ZIP: 22202  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: PatentIn Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/809,103b  
 : FILING DATE: 17-MAR-1997  
 : CLASSIFICATION: 800  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: FR 94.11040  
 : FILING DATE: 15-SEP-1994  
 : APPLICATION TYPE: WO PCT/FR95/01192  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: PATCH, Andrew J.  
 : REGISTRATION NUMBER: 32,925  
 : REFERENCE/DOCKET NUMBER: USB94AL CNR TOM  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: (703) 521-2297  
 : TELEX: 248425 EMBON  
 : INFORMATION FOR SEQ ID NO: 6:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 359 amino acids  
 : TYPE: amino acid  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: protein  
 : US-08-809-103b-6

Query Match 60.7%; Score 221; DB 3; Length 359;  
 Best Local Similarity 56.5%; Pred. No. 2,2e-21;  
 Matches 39; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

QY 2 LVNGFQVGRSARGCQTSNDLLLEALNASKEALQIIREKIPKYLQFPHNLSNLD 61  
 Db 111 LVNGFQVGRSARGCQTSNDLLLEALNASKEALQIIREKIPKYLQFPHNLSNLD 170  
 QY 62 RIFDKTPEP 70  
 Db 171 RVFQVPPAP 179

## RESULT 4

US-08-809-103b-8  
 : Sequence 8, Application US/08809103b  
 : Patent No. 6133505  
 : GENERAL INFORMATION:  
 : APPLICANT: GRONENBORN, Bruno  
 : TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT  
 : TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME  
 : NUMBER OF SEQUENCES: 17  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: YOUNG & THOMPSON  
 : STREET: 745 South 23rd Street  
 : CITY: Arlington  
 : STATE: Virginia

```

1  COUNTRY: U.S.A.
2  ZIP: 22202
3  COMPUTER READABLE FORM:
4  MEDIUM TYPE: FLOPPY disk
5  OPERATING SYSTEM: PC-DOS/MS-DOS
6  CURRENT APPLICATION DATA:
7  APPLICATION NUMBER: US/08/809, 103B
8  FILING DATE: 17-MAR-1997
9  CLASSIFICATION: 800
10 APPLICATION DATA: FR 94.11040
11 FILING DATE: 15-SEP-1994
12 APPLICATION NUMBER:
13 PRIOR APPLICATION NUMBER: WO PCT/FR95/01192
14 FILING DATE: 15-SEP-1995
15 ATTORNEY/AGENT INFORMATION:
16 NAME: PATCH, Andrew J, 925
17 TELEPHONE: (703) 521-2297
18 REFERENCE/DOCKET NUMBER: US894AL CNR TOM
19 TELECOMMUNICATION INFORMATION:
20 TELEPHONE: (703) 521-2297
21 TELEFAX: (703) 685-0573
22 TELEX: 248425 EMBON
23 INFORMATION FOR SEQ ID NO: 8:
24 SEQUENCE CHARACTERISTICS:
25 TYPE: amino acids
26 TOPOLOGY: linear
27 MOLECULE TYPE: protein
28 US-08-809-103B-8
29
30 Query Match 50.7%: Score 221; DB 3; Length 359;
31 Identical Similarity 56.5%; Pred: NO; 2.2e-21;
32 Matches 33; Conservative 11; Mismatches 19; Indels 0; Gaps 0;
33
34 Qy 2 LWCSEFQVDRSGRGQTSNDLLLEALNASSKEALQIIREKIPEKYLFOFHNL 61
35 Db 111 LLEWTFQIDGRSGRGQQTANDAYAKINAGSKQALQVKEIAPRDYVLFHFINSLD 170
36
37 Qy 62 RIFQKTPPEP 70
38 Db 171 KVFQVPPAP 179
39
40 RESULT 5
41 US-08-419-652-6
42 Sequence 6, Application US/08419652
43 Patent No. 5831007
44 GENERAL INFORMATION:
45 APPLICANT: Chua, Anne O
46 TITLE OF INVENTION: INTERLEUKIN-12 RECEPTOR
47 NUMBER OF SEQUENCES: 7
48 CORRESPONDENCE ADDRESS:
49 ADDRESSES: Hoffmann-La Roche Inc.
50 STREET: Nutley 40 Kingsland Street
51 CITY: Nutley, NJ 07110
52 STATE: New Jersey
53 COUNTRY: United States of America
54 ZIP: 07110-1199
55 COMPUTER READABLE FORM:
56 MEDIUM TYPE: Floppy disk
57 COMPUTER: PC compatible
58 OPERATING SYSTEM: MS-DOS
59 CURRENT APPLICATION DATA:
60 APPLICATION NUMBER: US/08/419,652
61 FILING DATE: 11-APR-1995
62 CLASSIFICATION: 530
63 PRIOR APPLICATION DATA:

```

```

1  APPLICATION NUMBER: US 08/248,532
2  FILING DATE: 31-MAY-1994
3  PRIOR APPLICATION DATA:
4  APPLICATION NUMBER: US 08/094,713
5  FILING DATE: 19-JUL-1993
6  ATTORNEY/AGENT INFORMATION:
7  NAME: Kass, Alan P.
8  REGISTRATION NUMBER: 32142
9  REFERENCE/DOCKET NUMBER: CD 9174
10 TELECOMMUNICATION INFORMATION:
11 TELEPHONE: (201) 235-4205
12 TELEFAX: (201) 235-3500
13 INFORMATION FOR SEQ ID NO: 6:
14 SEQUENCE CHARACTERISTICS:
15 LENGTH: 502 amino acids
16 TYPE: amino acid
17 TOPOLOGY: linear
18 MOLECULE TYPE: protein
19 FEATURE:
20 NAME/KEY: Region
21 FUNCTION: CD 9174
22 OTHER INFORMATION: /note= *Represents residues 98 to
23 OTHER INFORMATION: 731 of human granulocyte colony-stimulating
24 OTHER INFORMATION: factor-receptor.*
25 US-08-419-652-6
26
27 Query Match 16.8%: Score 61; DB 2; Length 602;
28 Identical Similarity 26.0%; Pred: NO; 7.4;
29 Matches 13; Conservative 9; Mismatches 22; Indels 1;
30
31 Qy 7 FOVDGRSGRGQTSNDLLLEALNASSKEALQIIREKIPEKYLFOFHNL 56
32 Db 37 FTLASEKRGNCQTGDSILDVCPKDGSHCC-----IPRKLHLLYQNM 80
33
34 RESULT 6
35 US-07-923-976-6
36 Sequence 6, Application US/07923976
37 Patent No. 5574136
38 GENERAL INFORMATION:
39 APPLICANT: Nagata, Shigekazu
40 APPLICANT: Fukunaga, Rikio
41 TITLE OF INVENTION: DNA Encoding Granulocyte
42 TYPE OF INVENTION: Colony-Stimulating Factor Receptor
43 NUMBER OF SEQUENCES: 8
44 CORRESPONDENCE ADDRESS:
45 ADDRESSES: Jones, Tullar & Cooper, P.C.
46 STREET: P.O. Box 2266 Eads Station
47 CITY: Arlington
48 STATE: Virginia
49 COUNTRY: United States of America
50 ZIP: 22202
51 COMPUTER READABLE FORM:
52 MEDIUM TYPE: Floppy disk
53 COMPUTER: IBM PC compatible
54 OPERATING SYSTEM: PC-DOS/MS-DOS
55 SOFTWARE: Patent In Release #1.0, Version #1.25
56 CURRENT APPLICATION DATA:
57 APPLICATION NUMBER: US/07/923,976
58 FILING DATE: 1992/922
59 CLASSIFICATION: 430
60 PRIOR APPLICATION DATA:
61 PRIOR APPLICATION NUMBER: JP 74539/1990
62 FILING DATE: 23-MAR-1990
63 PRIOR APPLICATION DATA:
64 APPLICATION NUMBER: JP 176629/1990
65 FILING DATE: 03-JUL-1990
66 PRIOR APPLICATION DATA:
67 APPLICATION NUMBER: PCT/JP91/00375
68 FILING DATE: 22-MAR-1991
69 ATTORNEY/AGENT INFORMATION:
70 NAME: Hellwege, James W.
71 REGISTRATION NUMBER: 28,808

```

```

; REFERENCE/DOCKET NUMBER: 514853
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-415-1500
; COMMERCIAL RELEASE #1.0, Version #1.25
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-923-976-6

```

```

Query Match 16.8%; Score 61; DB 1; Length 771;
Best Local Similarity 26.0%; Pred. No. 10;
Matches 13; Conservative 9; Mismatches 22; Indels 6; Gaps 1;

```

```

OY 7 FOVDCRGAGCGCOTSDNLLLEALNASKKEALQIRKIPKYLQPHNL 56
DB 156 FTLSFKSRGNCOTQGDSDILDCVPKDGSHCC-----IPRKHLLLYQNM 199

```

```

RESULT 7
; Patent No. 5422248
; APPLICANT: SMITH, CRAIG A.; LARSEN, ALF D.; SIMS, JOHN E.;
; TITLE OF INVENTION: DNA SEQUENCES ENCODING GRANULOCYTE-COLONY
; STIMULATING FACTOR RECEPTORS
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; FILING DATE: 15-JAN-1993
; PRIOR APPLICATION DATA: US/08/6.183
; FILING DATE: 15-JAN-1993
; FILING DATE: 24-SEP-1990
; APPLICATION NUMBER: 522,329
; APPLICATION NUMBER: 522,952
; FILING DATE: 03-APR-1990
; APPLICATION NUMBER: 416,306
; FILING DATE: 03-OCT-1989
; APPLICATION NUMBER: 412,816
; FILING DATE: 26-SEP-1989
; SEQ ID LENGTH: 783
; US-07-923-976-4
5422248-2

```

```

Query Match 16.8%; Score 61; DB 5; Length 783;
Best Local Similarity 26.0%; Pred. No. 11;
Matches 13; Conservative 9; Mismatches 22; Indels 6; Gaps 1;

```

```

OY 7 FOVDCRGAGCGCOTSDNLLLEALNASKKEALQIRKIPKYLQPHNL 56
DB 156 FTLSFKSRGNCOTQGDSDILDCVPKDGSHCC-----IPRKHLLLYQNM 199

```

```

RESULT 8
; US-07-923-976-4
; Sequence 8, Application US/07923976
; Patent No. 5574136
; GENERAL INFORMATION:
; APPLICANT: Nagata, Shigekazu
; APPLICANT: Fukunaga, Rikiko
; TITLE OF INVENTION: DNA Encoding Granulocyte
; COLONY-Stimulating Factor Receptor
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones, Tullar & Cooper, P.C.
; STREET: P.O. Box 2266 Eads Station
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/923.976
; FILING DATE: 19900222
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 74539/1990
; FILING DATE: 23-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 176629/1990
; FILING DATE: 03-JUL-1990
; .COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/923.976
; FILING DATE: 19920922
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 74539/1990
; FILING DATE: 23-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 176629/1990
; FILING DATE: 03-JUL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP91/00375
; FILING DATE: 22-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hellwege, James W.
; REFERENCE/DOCKET NUMBER: 514853
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-415-1500
; TELEFAX: 703-415-1508
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 836 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-923-976-4

```

```

Query Match 16.8%; Score 61; DB 1; Length 836;
Best Local Similarity 26.0%; Pred. No. 12;
Matches 13; Conservative 9; Mismatches 22; Indels 6; Gaps 1;

```

```

OY 7 FOVDCRGAGCGCOTSDNLLLEALNASKKEALQIRKIPKYLQPHNL 56
DB 156 FTLSFKSRGNCOTQGDSDILDCVPKDGSHCC-----IPRKHLLLYQNM 199

```

```

RESULT 9
; US-07-923-976-8
; Sequence 8, Application US/07923976
; Patent No. 5574136
; GENERAL INFORMATION:
; APPLICANT: Nagata, Shigekazu
; APPLICANT: Fukunaga, Rikiko
; TITLE OF INVENTION: DNA Encoding Granulocyte
; COLONY-Stimulating Factor Receptor
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones, Tullar & Cooper, P.C.
; STREET: P.O. Box 2266 Eads Station
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/923.976
; FILING DATE: 19900222
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 74539/1990
; FILING DATE: 23-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 176629/1990
; FILING DATE: 03-JUL-1990

```



PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/JF91/00375  
 FILING DATE: 22-MAR-1991  
 ANNUAL FEE INFORMATION:  
 NAME: Filing Agent: 28, 808  
 REGISTRATION NUMBER: 514853  
 REFERENCE/DOCKET NUMBER: 514853  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 703-415-1500  
 TELEFAX: 703-415-1508  
 INFORMATION FOR SEQ ID NO: 8:  
 SEQUENCE LENGTH: 863  
 LENGTH: 863 amino acids  
 TYPE: AMINO ACID  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-07-923-976-8

Query Match 16.8%; Score 61; DB 1; Length 863;  
 Best Local Similarity 26.0%; Pred. No. 12;  
 Matches 13; Conservative 9; Mismatches 22; Indels 1; Gaps 1;

Qy 7 FOVDGSRGCGTSDNDLLALNASSKEEALQIREKIPYLFQPHL 56  
 Db 156 FTLSKFKSGNGCTGDSILDCVFKDQSHCS-----IPKHLIIYQNM 199

RESULT 10  
 US-07-745-206A-2  
 Sequence 2, Application US/07745206A  
 Patent No. 5429921

GENERAL INFORMATION:  
 APPLICANT: Harpold, Michael  
 APPLICANT: Williams, Steven  
 APPLICANT: Williams, Mark  
 APPLICANT: McCue, Ann  
 APPLICANT: Feldman, Daniel  
 TITLE OF INVENTION: Human Calcium Channel Compositions and  
 NUMBER OF SEQUENCES: 32  
 CORRESPONDENCE ADDRESS:  
 STREET: 135 S. LaSalle  
 CITY: Chicago  
 STATE: Illinois  
 COUNTRY: U.S.A.  
 ZIP: 60603

COMPUTER READABLE FORM:  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 FILING DATE: 19910815  
 FILING DATE: 20-FEB-1990

CLASSIFICATION: 435  
 NAME: Filing Agent: 435  
 REGISTRATION NUMBER: 51504  
 REFERENCE/DOCKET NUMBER: 51504  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 312-372-7842  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2161 amino acids  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-07-745-206A-2

Query Match 16.1%; Score 58.5; DB 1; Length 2161;  
 Best Local Similarity 28.4%; Pred. No. 88;

Matches 21; Conservative 6; Mismatches 18; Indels 29; Gaps 2;  
 Qy 4 WEPEVDGSRGARGCGTSDNDLL-----NDLL-----FALNASSK 34  
 Db 707 WNAVDDGDMATGPRSSGNCVITFTLTCGNTYLLNVLATANDNLADRESLTKAKR 766  
 Qy 35 BEALQIREKIPK 48  
 Db 767 BEAEKEREKKIARK 780

RESULT 11  
 US-08-455-543A-49  
 Sequence 49, Application US/08455543A  
 Patent No. 5792846  
 GENERAL INFORMATION:  
 APPLICANT: Harpold, Michael  
 APPLICANT: Ellis, Steven  
 APPLICANT: Williams, Mark  
 APPLICANT: McCue, Ann  
 APPLICANT: Feldman, Daniel  
 APPLICANT: Brenner, Robert  
 TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND  
 NUMBER OF SEQUENCES: 57  
 CORRESPONDENCE ADDRESS:  
 STREET: 166 Union Street  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 92101-2926

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: PC-DOS/MS-DOS Version 1.5  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/455,543A  
 FILING DATE: May 31, 1995

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/223,305  
 FILING DATE: April 4, 1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 07/868,354  
 FILING DATE: April 10, 1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/745,206  
 FILING DATE: 15-AUG-1991

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/603,751  
 FILING DATE: 04-APR-1989  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: WO PCT/US89/01408  
 FILING DATE: 04-APR-1989  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/176,899  
 FILING DATE: 04-APR-1988

ATTORNEY/AGENT INFORMATION:  
 NAME: Seidman, Stephanie L.  
 REGISTRATION NUMBER: 5719  
 REFERENCE/DOCKET NUMBER: 6362-52517  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (619)238-0999  
 TELEFAX: (619)238-0062  
 INFORMATION FOR SEQ ID NO: 49:  
 SEQUENCE CHARACTERISTICS:

? LENGTH: 2161 amino acids  
? TYPE: amino acid  
? STRANDEDNESS: single  
? MOLECULE TYPE: linear  
? FRAGMENT TYPE: internal  
US-08-455-543A-49

Query Match 16.1%; Score 58.5; DB 1; Length 2161;  
Best Local Similarity 28.4%; Pred. No. 88; Mismatches 18; Indels 29; Gaps 2;  
Matches 21; Conservative 6;

QY 4 WGEFQVQGSARGCQCTS-----NDLL-----EALNASSK 34  
DB 707 WNAVWYDGMVYCGSSGMVCIYFILFCGNYILLVFLAVNDLADAEISLNTAK 766  
QY 35 BEALQIIREKPEK 48  
DB 767 EAEERKKKIARK 780

## RESULT 12

US-08-455-543A-51  
? Sequence 51, Application US/08455543A  
? Patent No. 57952846  
? GENERAL INFORMATION:  
? APPLICANT: Harpold, Michael  
? APPLICANT: Ellis, Steven  
? APPLICANT: Williams, Mark  
? APPLICANT: Feldman, Daniel  
? APPLICANT: McCue, Ann  
? APPLICANT: Brenner, Robert  
? TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND  
? TITLE OF INVENTION: METHODS  
? NUMBER OF SEQUENCES: 57  
? CORRESPONDENCE ADDRESS:  
? ADDRESSEE: Brown, Martin, Haller & McClain  
? STREET: 1660 Union Street  
? CITY: San Diego  
? STATE: California  
? COUNTRY: USA  
? ZIP: 92101-2926

COMPUTER READABLE FORM:  
? MEDIUM TYPE: Diskette  
? OPERATING SYSTEM: DOS  
? SOFTWARE: FASTSEQ Version 1.5  
? CURRENT APPLICATION DATA:  
? APPLICATION NUMBER: US/08/455-543A  
? FILING DATE: April 4, 1994  
? PRIOR APPLICATION DATA:  
? APPLICATION NUMBER: 08/223,305  
? FILING DATE: April 4, 1994  
? PRIOR APPLICATION DATA:  
? APPLICATION NUMBER: 07/868,354  
? FILING DATE: April 10, 1992  
? PRIOR APPLICATION DATA: US 07/745,206  
? FILING DATE: 15-AUG-1991  
? PRIOR APPLICATION DATA:  
? APPLICATION NUMBER: US 07/620,250  
? FILING DATE: 30-NOV-1990  
? PRIOR APPLICATION DATA:  
? APPLICATION NUMBER: US 07/482,384  
? FILING DATE: 04-FEB-1990  
? PRIOR APPLICATION DATA:  
? APPLICATION NUMBER: US 07/603,751  
? FILING DATE: 04-APR-1989  
? PRIOR APPLICATION DATA:  
? APPLICATION NUMBER: WO PCT/US89/01408  
? FILING DATE: 04-APR-1989  
? PRIOR APPLICATION DATA:

? APPLICATION NUMBER: US 07/176,899  
? FILING DATE: 04-APR-1988  
? ATTORNEY/AGENT INFORMATION:  
? NAME: Stephen L. Sciamanna  
? REGISTRATION NUMBER: 1779  
? REFERENCE/DOCKET NUMBER: 6362-52517  
? TELECOMMUNICATION INFORMATION:  
? TELEPHONE: (619)238-0999  
? TELEFAX: (619)238-0062  
? INFORMATION FOR SEQ ID NO: 51:  
? SEQUENCE CHARACTERISTICS:  
? LENGTH: 2161 amino acids  
? STRANDEDNESS: single  
? MOLECULE TYPE: protein  
? FRAGMENT TYPE: internal  
US-08-455-543A-51

Query Match 16.1%; Score 58.5; DB 1; Length 2161;  
Best Local Similarity 28.4%; Pred. No. 88; Mismatches 18; Indels 29; Gaps 2;  
Matches 21; Conservative 6;

QY 4 WGEFQVQGSARGCQCTS-----NDLL-----EALNASSK 34  
DB 707 WNAVWYDGMVYCGSSGMVCIYFILFCGNYILLVFLAVNDLADAEISLNTAK 766  
QY 35 BEALQIIREKPEK 48  
DB 767 EAEERKKKIARK 780

## RESULT 13

US-08-223-305C-49  
? Sequence 49, Application US/08223305C  
? Patent No. 5651824  
? GENERAL INFORMATION:  
? APPLICANT: Harpold, Michael  
? APPLICANT: Ellis, Steven  
? APPLICANT: Williams, Mark  
? APPLICANT: Feldman, Daniel  
? APPLICANT: McCue, Ann  
? APPLICANT: Brenner, Robert  
? TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND  
? TITLE OF INVENTION: METHODS  
? NUMBER OF SEQUENCES: 57  
? CORRESPONDENCE ADDRESS:  
? ADDRESSEE: Brown, Martin, Haller & McClain  
? STREET: 1660 Union Street  
? CITY: San Diego  
? STATE: California  
? COUNTRY: USA  
? ZIP: 92101-2926

COMPUTER READABLE FORM:  
? MEDIUM TYPE: Diskette  
? COMPUTER: IBM Compatible  
? OPERATING SYSTEM: DOS  
? SOFTWARE: FASTSEQ Version 1.5  
? CURRENT APPLICATION DATA:  
? APPLICATION NUMBER: US/08/223,305C  
? FILING DATE: April 4, 1994  
? PRIOR APPLICATION DATA:  
? APPLICATION NUMBER: 07/868,354  
? FILING DATE: April 10, 1992  
? PRIOR APPLICATION DATA: US 07/745,206  
? FILING DATE: 15-AUG-1991  
? PRIOR APPLICATION DATA:  
? APPLICATION NUMBER: US 07/620,250  
? FILING DATE: 30-NOV-1990  
? PRIOR APPLICATION DATA:  
? APPLICATION NUMBER: US 07/482,384

```

; FILING DATE: 20-FEB-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/603,751
; FILING DATE: 15-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: NO PCT/US89/01408
; FILING DATE: 04-APR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/482,384
; FILING DATE: 20-FEB-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/603,751
; FILING DATE: 04-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: NO PCT/US89/01408
; FILING DATE: 04-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/176,999
; FILING DATE: 04-APR-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Stephen L.
; REGISTRATION NUMBER: 31779
; REFERENCE/DOCKET NUMBER: 52516 (P519739)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)238-0999
; TELEFAX: (619)238-0062
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2161
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: Internal
US-08-223-305C-49

```

```

Query Match 16.1%; Score 58.5; DB 2; Length 2161;
Best Local Similarity 28.4%; Pred. No. 88;
Matches 21; Conservative 6; Mismatches 18; Indels 29; Gaps 2;

```

```

Qy 4 WEEFQVDRSGRAGCGQTS-----NDLLL-----BALNASK 34
Db 707 WNAVWDGIMAYGGPSSSNVICVIFILPFCGNYILNFWLAVNVLADSLNIAK 766
Qy 35 EEAQLIIREKIPEK 48
Db 767 EBAEKERKKIARK 780

```

```

RESULT 14
US-08-223-305C-51
; Sequence 51: Application US/08223305C
; Patent No. 5851824
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feldman, Daniel
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/223,305C
; FILING DATE: April 4, 1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/868,354
; FILING DATE: April 10, 1992

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/745,206
; FILING DATE: 15-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/620,250
; FILING DATE: 30-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/482,384
; FILING DATE: 20-FEB-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/603,751
; FILING DATE: 04-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: NO PCT/US89/01408
; FILING DATE: 04-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/176,999
; FILING DATE: 04-APR-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Stephen L.
; REGISTRATION NUMBER: 31779
; REFERENCE/DOCKET NUMBER: 52516 (P519739)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)238-0999
; TELEFAX: (619)238-0062
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2161
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: Internal
US-08-223-305C-51

```

```

Query Match 16.1%; Score 58.5; DB 2; Length 2161;
Best Local Similarity 28.4%; Pred. No. 88;
Matches 21; Conservative 6; Mismatches 18; Indels 29; Gaps 2;

```

```

Qy 4 WEEFQVDRSGRAGCGQTS-----NDLLL-----BALNASK 34
Db 707 WNAVWDGIMAYGGPSSSNVICVIFILPFCGNYILNFWLAVNVLADSLNIAK 766
Qy 35 EEAQLIIREKIPEK 48
Db 767 EBAEKERKKIARK 780

```

```

RESULT 15
US-08-313-393-2
; Sequence 47: Application US/08311363
; Patent No. 5876958
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feldman, Daniel
; APPLICANT: McCue, Anne
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: Human Calcium Channel Compositions and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatenLin Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/311,363
/ PENDING APPLICATION DATA:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/745,206
/ FILING DATE: 15-AUG-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Seidman, Stephanie L.
/ REGISTRATION NUMBER: 33,779
/ REFERENCE/DOCKET NUMBER: 6362-51506
/ TELEPHONE: (619)238-0999
/ TELEPHONE: (619)238-0999
/ TELEFAX: (619)238-0062
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2161 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-311-363-2

Query Match 15.1%; Score 58.5; DB 2; Length 2161;
Best Local Similarity 28.4%; Pred. No. 88;
Matches 21; Conservative 6; Mismatches 18; Indels 29; Gaps 2:

QY 4 WGEFQVDSRGSGCOTS-----NDLLL-----EALNASSK 34
      ||| ||| ||| |||
DB 707 WNAVWYDGMATGPGSSGMVCIYFIILTCGNVILLNVLAVDNLADAESINTAQK 766
      ||| ||| ||| |||

QY 35 EEALQIIREKIPEK 48
      ||| : ||| |
DB 767 EEAEKERKRIARK 780
```

Search completed: February 3, 2001, 02:19:52  
Job time: 1714 sec



XX Example 5; Page 100-102; 132pp; English.

XX This sequence comprises the bean golden mosaic virus (BGMV) CI  
 CC protein that is required for replication. The invention involves  
 CC production of transgenic plants containing DNA comprising CI or AC1  
 CC wild-type or mutant sequences that negatively interfere in trans  
 CC with geminiviral replication during infection. Such transgenic  
 CC plants are resistant to viral infection. The AC1/CI genes are  
 CC especially from BGWV, tomato mottle virus or tomato yellow leaf  
 CC curl virus (see T93282-93) and encode polypeptides (see W34324-35)  
 CC and/or the NTP-binding domains.

XX Sequence 353 AA;

Query Match 61.3%; Score 223; DB 18; Length 353;

Best Local Similarity 60.0%; Pred. No. 2.4e-20;

Matches 42; Conservative 12; Mismatches 10; Indels 0; Gaps 0;

QY 1 TLVNGEFQVDRSGRGCGTSDNLLLEALNASSKEALQIREKIPKYLFOFHNLSNL 60

Db 110 TLWGGFQVDRSGRGCGTSDNLLLEALNASSKEALQIREKIPKYLFOFHNLSNL 169

QY 61 DRIPKTPPEP 70

Db 170 GRIFVKYPEP 179

RESULT 2

W34332

ID W34332 standard; Protein; 353 AA.

XX W34332;

XX W34332;

DT 27-APR-1998 (first entry)

XX Bean golden mosaic geminivirus C1 BGA190 mutant.

XX Geminivirus; BGWV; CI gene; transdominant mutation;  
 XX transgenic plant; disease resistance.

XX Bean golden mosaic virus type II isolate Guatemala.

XX W09739110-A1.

XX 23-OCT-1997.

XX 15-APR-1997; 97WO-US06300.

XX 16-APR-1996; 96US-0015517.

PA (SEMI-) SEMINIS VEGETABLE SEEDS INC.

PA (WISC) WISCONSIN ALUMNI RES FOUND.

XX Ahlquist PG, Hanson SF, Lau HT, Maxwell DP, Stout JT;

XX WPI. 1997-526447/48.

XX N-PSDB; T93290.

XX Transgenic plants expressing geminivirus AC1 and CI wild-type and

XX mutant genes - have increased resistance to geminivirus infection  
 XX e.g. tomato mottle virus, tomato yellow leaf curl virus or bean  
 XX golden mosaic geminivirus

XX Example 5; Page 103-105; 132pp; English.

XX This protein comprises a control mutant of the bean golden mosaic  
 CC virus (BGMV) CI protein (see W34338) that is required for  
 CC replication. It is encoded by mutated CI open reading frame  
 CC BGA1390 (see T93290). The invention involves production of  
 CC transgenic plants containing DNA comprising geminivirus CI or AC1

XX

CC wild-type or mutant sequences that negatively interfere in trans  
 CC with geminiviral replication during infection. Such transgenic  
 CC plants are resistant to viral infection. The AC1/CI genes are  
 CC especially from BGWV, tomato mottle virus or tomato yellow leaf  
 CC curl virus (see T93282-93) and encode polypeptides (see W34324-35)  
 CC that have mutations in the highly conserved DNA-nicking and/or the  
 CC NTP-binding domains.

XX Sequence 353 AA;

Query Match 61.3%; Score 223; DB 18; Length 353;

Best Local Similarity 60.0%; Pred. No. 2.4e-20;

Matches 42; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLVNGEFQVDRSGRGCGTSDNLLLEALNASSKEALQIREKIPKYLFOFHNLSNL 60

Db 110 TLWGGFQVDRSGRGCGTSDNLLLEALNASSKEALQIREKIPKYLFOFHNLSNL 169

QY 61 DRIPKTPPEP 70

Db 170 GRIFVKYPEP 179

RESULT 3

W34333

ID W34333 standard; Protein; 353 AA.

XX W34333;

XX W34333;

DT 27-APR-1998 (first entry)

XX Bean golden mosaic geminivirus C1 BGA221 mutant.

XX Geminivirus; BGWV; CI gene; transdominant mutation;  
 XX transgenic plant; disease resistance.

XX Bean golden mosaic virus type II isolate Guatemala.

XX W09739110-A1.

XX 23-OCT-1997.

XX 15-APR-1997; 97WO-US06300.

XX 16-APR-1996; 96US-0015517.

PA (SEMI-) SEMINIS VEGETABLE SEEDS INC.

PA (WISC) WISCONSIN ALUMNI RES FOUND.

XX Ahlquist PG, Hanson SF, Lau HT, Maxwell DP, Stout JT;

XX WPI. 1997-526447/48.

XX N-PSDB; T93291.

XX Transgenic plants expressing geminivirus AC1 and CI wild-type and  
 XX mutant genes - have increased resistance to geminivirus infection  
 XX e.g. tomato mottle virus, tomato yellow leaf curl virus or bean  
 XX golden mosaic geminivirus

XX Example 5; Page 107-109; 132pp; English.

XX This protein comprises a transdominant lethal mutant of the bean  
 CC golden mosaic virus (BGMV) CI protein (see W34338) that is required  
 CC for replication. It is encoded by mutation CI open reading frame  
 CC BGA221 (see T93291) and carries a mutation in the NTP-binding  
 CC domain. The invention involves production of transgenic plants  
 CC containing DNA comprising geminivirus CI or AC1 wild-type or mutant  
 CC sequences that negatively interfere in trans with geminiviral  
 CC replication during infection. Such transgenic plants are resistant  
 CC to viral infection. The AC1/CI genes are especially from BGWV,  
 CC tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)  
 CC and encode polypeptides (see W34324-35) that have mutations in the

Query Match 61.3% Score 223 DB 18: Length 353:

0.

Qy 1 TLVWGEFQVDGRSARGGCQTNDLLLEALNASKEEALQIREKIPKEYLQFPHNLNSNL 60  
 :  
pb 110 tiwvcafvdarsargaasandsvakalnadsjesaltllkeegakdvvgahnrnsnl 169

```

QY 61 DRIFDKTPEP 70
Db 170 erifkvppep 179

RESULT 6
ID P70407 standard; Protein: 353 AA.
AC P70407;
XX
XX 02-MAY-1991 (first entry)
DE ORF 4 gene product of Bean Golden Mosaic virus.
XX
XX Plant vector.
XX
XX Bean golden mosaic virus.
OS
XX JP61257186-A.
XX
XX 14-NOV-1986.
XX
XX 10-MAY-1985; 85JP-0098108.
XX
XX 10-MAY-1985; 85JP-0098108.
XX
XX (TEIJ ) TEIJIN KK.
XX
XX WPI: 1987-159662/23.
DR N-PSDB; N70630.
XX
XX New DNA and hybrid DNA - used for recombinant vector of plants.
XX
XX Disclosure; Fig 6; 24pp; Japanese.
XX
XX The sequence encoding this protein may be taken along with the -a
CC DNA sequence and a suitable resistance gene, and used to create a
CC recombinant plant vector.
CC See also N70629.
XX
XX Sequence 353 AA:

Query Match 60.7%; Score 221; DB 8; Length 353;
Best Local Similarity 60.0%; Pred. No. 4, 3e-20;
Matches 42; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLWGSEFQVDSRGCGTSNDLLLEALNASKEALQITRETPKYLFOFHNLNML 60
Db 110 lfwgsgfvdgrsggqtdsdayakalnadsalsltlkeepkoyvignmrlnl 169

QY 61 DRIFDKTPEP 70
Db 170 erifkvppep 179

RESULT 7
ID R88870 standard; Protein: 359 AA.
AC R88870;
XX
XX 07-NOV-1996 (first entry)
XX
XX Sardinian tomato yellow leaf curl virus mutated Rep protein (K227A).
XX
XX Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
XX modification; mutation; viral replication; deficient; inhibition;
XX viral resistance; geminivirus; tomato yellow leaf curl virus;
XX Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
XX All protein; dominant negative phenotype.
XX
XX Sardinian tomato yellow leaf curl virus.
XX
XX Key Location/Qualifiers

QY 61 DRIFDKTPEP 70
Db 170 erifkvppep 179

RESULT 6
ID P70407 standard; Protein: 353 AA.
AC P70407;
XX
XX 02-MAY-1991 (first entry)
DE ORF 4 gene product of Bean Golden Mosaic virus.
XX
XX Plant vector.
XX
XX Bean golden mosaic virus.
OS
XX JP61257186-A.
XX
XX 14-NOV-1986.
XX
XX 10-MAY-1985; 85JP-0098108.
XX
XX 10-MAY-1985; 85JP-0098108.
XX
XX (TEIJ ) TEIJIN KK.
XX
XX WPI: 1987-159662/23.
DR N-PSDB; N70630.
XX
XX New DNA and hybrid DNA - used for recombinant vector of plants.
XX
XX Disclosure; Fig 6; 24pp; Japanese.
XX
XX The sequence encoding this protein may be taken along with the -a
CC DNA sequence and a suitable resistance gene, and used to create a
CC recombinant plant vector.
CC See also N70629.
XX
XX Sequence 353 AA:

Query Match 60.7%; Score 221; DB 8; Length 353;
Best Local Similarity 60.0%; Pred. No. 4, 3e-20;
Matches 42; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLWGSEFQVDSRGCGTSNDLLLEALNASKEALQITRETPKYLFOFHNLNML 60
Db 110 lfwgsgfvdgrsggqtdsdayakalnadsalsltlkeepkoyvignmrlnl 169

QY 61 DRIFDKTPEP 70
Db 170 erifkvppep 179

RESULT 7
ID R88870 standard; Protein: 359 AA.
AC R88870;
XX
XX 07-NOV-1996 (first entry)
XX
XX Sardinian tomato yellow leaf curl virus mutated Rep protein (K227H).
XX
XX Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
XX modification; mutation; viral replication; deficient; inhibition;
XX viral resistance; geminivirus; tomato yellow leaf curl virus;
XX Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
XX All protein.
XX
XX Sardinian tomato yellow leaf curl virus.
XX
XX Key Location/Qualifiers

```

```

XX OS Sardinian tomato yellow leaf curl virus.
XX
XX Key Location/Qualifiers
XX Misc-difference 227 /note= "wild-type Lys has been replaced by Ala"
XX
XX W09508573-AL.
XX
XX 21-MAR-1996.
XX
XX 15-SEP-1995; 95WO-FR01192.
XX
XX 15-SEP-1994; 94FR-0011040.
XX
XX (CNRS ) CENT NAT RECH SCI.
XX
XX Gronenborn B;
XX
XX WPI: 1996-179947/18.
XX
XX N-PSDB; T12904.
XX
XX Prodn. of virus-resistant transgenic plants - using mutated genomic
XX sequence from phytopathogenic DNA virus
XX
XX Disclosure; Fig 13; 93pp; French.
XX
XX Mutation of consensus amino acids in the NTP-binding site of
XX geminivirus rep proteins used to produce replication deficient
XX transgenic plants that are resistant to, or tolerant of, the native
XX virus. The present sequence is a mutant form of the Rep (or C1)
XX protein from the Sardinian isolate of tomato yellow leaf curl virus
XX (STYLCV) in which the wild-type Lys227 residue has been changed to an
XX Ala residue; transgenic Nicotiana benthamiana plants generated by
XX transformation with the mutated virus were found to be resistant to
XX STYLCV, i.e. the mutation results in a dominant negative phenotype.
XX
XX Sequence 359 AA:

Query Match 60.7%; Score 221; DB 17; Length 359;
Best Local Similarity 56.5%; Pred. No. 4, 4e-20;
Matches 35; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

QY 2 LWGSEFQVDSRGCGTSNDLLLEALNASKEALQITRETPKYLFOFHNLNML 61
Db 111 lfwgsgfvdgrsggqtdsdayakalnagsqaldvikelaprdyvlthhainsld 170

QY 62 DRIFDKTPEP 70
Db 171 kvifkvppep 179

RESULT 8
ID R88871 standard; Protein: 359 AA.
XX
XX R88871;
XX
XX 07-NOV-1996 (first entry)
XX
XX Sardinian tomato yellow leaf curl virus mutated Rep protein (K227H).
XX
XX Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
XX modification; mutation; viral replication; deficient; inhibition;
XX viral resistance; geminivirus; tomato yellow leaf curl virus;
XX Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
XX All protein.
XX
XX Sardinian tomato yellow leaf curl virus.
XX
XX Key Location/Qualifiers

```



```

FT  Misc-difference 227
XX  /note- "wild-type lys has been replaced by His"
XX  W09608573-A1.
XX  21-MAR-1996.
PD  15-SEP-1995; 95WO-FR01192.
XX  15-SEP-1994; 94FR-0011040.
XX  (CNRS ) CENT NAT RECH SCI.
XX  Gronenborn B;
PI  WPI: 1996-179947/18.
XX  N-PSDB; T12905.
XX  Prodn. of virus-resistant transgenic plants - using mutated genomic
XX  sequence from phytopathogenic DNA virus
XX  Disclosure; Fig 13; 93pp; French.
XX  Mutation of consensus amino acids in the NTP-binding site of
XX  geminivirus Rep protein is used to produce replication deficient
XX  viruses. The mutated viral nucleic acid is used for producing
XX  transgenic plants that are resistant to, or tolerant of, the native
XX  virus. The present sequence is a mutant form of the Rep (or CI)
XX  protein from the Sardinian isolate of tomato yellow leaf curl virus
XX  (STYLCV) in which the wild-type Lys227 residue has been changed to
XX  His residue; transgenic Nicotiana benthamiana plants generated by
XX  transformation with the mutated virus were not resistant to STYLCV.
XX  In contrast, plants transformed with a virus in which Lys227 had been
XX  replaced by Ala were found to be resistant.
XX  Sequence 359 AA:

Query Match 60.7%; Score 221; DB 17; Length 359;
Best Local Similarity 56.5%; Pred. No. 4.4e-20;
Matches 39; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

Oy  2 LWMGEFQVDSRGCGTSDNLLLEALNASSKEALQITREKIPKYLFOFINHNSLD 61
Db  111 lewtqtdqrsrggqtdandayakainagskqldvikelaprdyvlhfninsld 170

Oy  62 RIFDKTPPE 70
Db  171 kvfqvppap 179

RESULT 9
ID R88872 standard; Protein; 359 AA.
AC R88872;
XX 07-NOV-1996 (first entry)
XX Sardinian tomato yellow leaf curl virus mutated Rep protein (K227R).
XX Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
XX modification; mutation; viral replication; deficit; inhibition;
XX viral resistance; geminivirus; tomato yellow leaf curl virus;
XX Sardinian isolate; STYLCV; transgenic plant; P-loop; CI protein;
XX ALI protein.
XX Sardinian tomato yellow leaf curl virus.
XX key Location/Qualifiers
XX Misc-difference 227
XX /note- "wild-type lys has been replaced by Arg"

```

```

PN  W09608573-A1.
PD  21-MAR-1996.
XX  15-SEP-1995; 95WO-FR01192.
XX  15-SEP-1994; 94FR-0011040.
XX  (CNRS ) CENT NAT RECH SCI.
XX  Gronenborn B;
PI  WPI: 1996-179947/18.
XX  N-PSDB; T12906.
XX  Prodn. of virus-resistant transgenic plants - using mutated genomic
XX  sequence from phytopathogenic DNA virus
XX  Disclosure; Fig 13; 93pp; French.
XX  Mutation of consensus amino acids in the NTP-binding site of
XX  geminivirus Rep protein is used to produce replication deficient
XX  viruses. The mutated viral nucleic acid is used for producing
XX  transgenic plants that are resistant to, or tolerant of, the native
XX  virus. The present sequence is a mutant form of the Rep (or CI)
XX  protein from the Sardinian isolate of tomato yellow leaf curl virus
XX  (STYLCV) in which the wild-type Lys227 residue has been changed to
XX  His residue; transgenic Nicotiana benthamiana plants generated by
XX  transformation with the mutated virus were not resistant to STYLCV.
XX  In contrast, plants transformed with a virus in which Lys227 had been
XX  replaced by Ala were found to be resistant.
XX  Sequence 359 AA:

Query Match 60.7%; Score 221; DB 17; Length 359;
Best Local Similarity 56.5%; Pred. No. 4.4e-20;
Matches 39; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

Oy  2 LWMGEFQVDSRGCGTSDNLLLEALNASSKEALQITREKIPKYLFOFINHNSLD 61
Db  111 lewtqtdqrsrggqtdandayakainagskqldvikelaprdyvlhfninsld 170

Oy  62 RIFDKTPPE 70
Db  171 kvfqvppap 179

RESULT 10
ID R88872 standard; Protein; 361 AA.
AC R88872;
XX 27-APR-1998 (first entry)
XX Tomato mottle virus AC1 protein.
XX Geminivirus; ToMoV; AC1 gene; transdominant mutation;
XX transgenic plant; disease resistance.
XX Tomato mottle virus isolate Florida.
XX W09739110-AL.
XX 23-OCT-1997.
XX 15-APR-1997; 97MO-US06300.
XX 16-APR-1996; 96OS-0015517.
XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.
XX (WISCI-) WISCONSIN ALUMNI RES FOUND.

```

XX PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;  
 XX DR WPI: 1997-526447/48.  
 XX DR N-PSDB: T93284.  
 XX PT Transgenic plants expressing geminivirus AC1 and CI wild-type and  
 XX PT mutant genes - have increased resistance to geminivirus infection  
 XX PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean  
 XX PT golden mosaic geminivirus  
 XX PS Example 3.4: Page 57-58; 132pp; English.  
 XX CC This protein comprises the wild-type AC1 protein of tomato mottle  
 CC virus (ToMoV), a geminivirus that has a bipartite genome. The AC1  
 CC gene (see T93284) must be expressed for efficient replication of  
 CC the two genomic components, DNA-A and DNA-B. The AC1 protein has a  
 CC DNA binding site specific to the DNA-A common region, a DNA nicking  
 CC activity, and an NTP binding activity. The invention involves  
 CC wild-type or mutant transgenic plants containing DNA comprising AC1 or CI  
 CC wild-type or mutant genes. The invention also involves transgenic  
 CC plants with geminiviral replication during infection. Such transgenic  
 CC plants are resistant to viral infection. The AC1/CI genes are  
 CC especially from ToMoV, tomato yellow leaf curl virus or bean golden  
 CC mosaic geminivirus (see T93282-93) and encode polypeptides (see  
 CC W34324-35) that have mutations in the highly conserved DNA-nicking  
 CC domain and/or the WPI-binding domain.  
 XX SQ Sequence 361 AA;

Query Match 60.7%; Score 221; DB 18; Length 361;  
 Best Local Similarity 57.1%; Pred. No. 4.4e-20;  
 Matches 40; Conservative 15; Mismatches 15; Indels 0; Gaps 0;  
 QY 1 TLVWGEFQVDRSGAGCGTSDNLLLEALNASSKEALQITREKIPKYLQFHNLSNL 60  
 DB 110 LLEWdfqldgrsrggqgaandsyakaInassvqsalavlrceqpkfviqunhirsnl 169  
 QY 61 DRIFDKTPPEP 70  
 DB 170 erifakepep 179  
 ||||| |||  
 ||||| |||

RESULT 11  
 W34324  
 ID W34324 standard; Protein; 361 AA.  
 XX W34324;  
 XX W34324;  
 DT 27-APR-1998 (first entry)  
 XX Tomato mottle virus AC1 mutant ToMoV-AC1dml.  
 XX Geminivirus; ToMoV-AC1dml: AC1 gene; transdominant mutation;  
 KW transgenic plant; disease resistance.  
 OS Tomato mottle virus isolate Florida.  
 OS Synthetic.  
 XX W09739110-A1.  
 PN W09739110-A1.  
 XX 23-OCT-1997.  
 PD 15-APR-1997; 97WO-US06300.  
 XX 15-APR-1997; 96US-0015517.  
 PR 16-APR-1996; 96US-0015517.  
 XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.  
 PA (WISC ) WISCONSIN ALUMNI RES FOUND.  
 XX Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;  
 XX WPI: 1997-526447/48.  
 XX DR N-PSDB: T93283.

DR WPI: 1997-526447/48.  
 DR N-PSDB: T93282.  
 XX Transgenic plants expressing geminivirus AC1 and CI wild-type and  
 XX PT mutant genes - have increased resistance to geminivirus infection  
 XX PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean  
 XX PT golden mosaic geminivirus  
 XX PS Example 3.4: Page 60-62; 132pp; English.  
 XX CC This protein comprises a transdominant lethal mutant, designated  
 CC ToMoV-AC1dml, of tomato mottle virus (ToMoV) AC1 protein (see  
 CC W34324) is encoded by a mutant AC1 gene (see T93282) of  
 CC ToMoV. The AC1 gene is a member of the AC1 gene family in its  
 CC NTP-binding domains. The AC1 gene (see also T93294) must be  
 CC expressed for efficient replication of the two genomic components,  
 CC DNA-A and DNA-B, of the bipartite ToMoV genome. The invention  
 CC involves production of transgenic plants containing DNA comprising  
 CC geminivirus AC1 or CI wild-type or mutant sequences that negatively  
 CC interfere in trans with geminiviral replication during infection.  
 CC Transgenic plants are resistant to viral infection. The  
 CC AC1/CI genes are especially from ToMoV, tomato yellow leaf curl  
 CC virus or bean golden mosaic geminivirus (see T93282-93) and encode  
 CC polypeptides (see W34324-35) that have mutations in the highly  
 CC conserved DNA-nicking and/or NTP-binding domains.  
 XX SQ Sequence 361 AA;

Query Match 60.7%; Score 221; DB 18; Length 361;  
 Best Local Similarity 57.1%; Pred. No. 4.4e-20;  
 Matches 40; Conservative 15; Mismatches 15; Indels 0; Gaps 0;  
 QY 1 TLVWGEFQVDRSGAGCGTSDNLLLEALNASSKEALQITREKIPKYLQFHNLSNL 60  
 DB 110 LLEWdfqldgrsrggqgaandsyakaInassvqsalavlrceqpkfviqunhirsnl 169  
 QY 61 DRIFDKTPPEP 70  
 DB 170 erifakepep 179  
 ||||| |||  
 ||||| |||

RESULT 12  
 W34325  
 ID W34325 standard; Protein; 361 AA.  
 XX W34325;  
 XX W34325;  
 DT 27-APR-1998 (first entry)  
 XX Tomato mottle virus AC1 mutant ToMoV-AC1dml.  
 XX Geminivirus; ToMoV-AC1dml: AC1 gene; transdominant mutation;  
 KW transgenic plant; disease resistance.  
 OS Tomato mottle virus isolate Florida.  
 OS Synthetic.  
 XX W09739110-A1.  
 PN W09739110-A1.  
 XX 23-OCT-1997.  
 PD 15-APR-1997; 97WO-US06300.  
 XX 15-APR-1997; 96US-0015517.  
 PR 16-APR-1996; 96US-0015517.  
 XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.  
 PA (WISC ) WISCONSIN ALUMNI RES FOUND.  
 XX Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;  
 XX WPI: 1997-526447/48.  
 XX DR N-PSDB: T93283.



Search completed: February 3, 2001, 02:15:30  
Job time: 3437 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:24:29 ; Search time 83.07 seconds  
(without alignments)  
27.213 Million cell updates/sec

Title: us-09-289-346a-9

Perfect score: 362

Sequence: 1 TLWGEFQVAGSARGSCOT.....FQPHNLSNLDRIKTPPEP 70

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	356	98.3	352	1 VAL1_TGMV	P03567 tomato gold
2	256	70.7	361	1 VAL1_PYMV	P27258 potato yell
3	243	67.1	358	1 VAL1_CLVK	P14982 cassava lat
4	243	67.1	358	1 VAL1_CLVK	P14982 cassava lat
5	236	65.2	362	1 VAL1_TYLCV	P36279 cassava yell
6	229	63.3	359	1 VAL1_TYLCV	P38609 tomato yell
7	225	62.2	349	1 VAL1_PHVU	P06923 pepper hus
8	222	61.3	359	1 VAL1_TYLCV	P06923 pepper hus
9	220	60.8	353	1 VAL1_BGMV	P27260 tomato yell
10	219	60.5	355	1 VAL1_BGMV	P05175 bean golden
11	216	59.7	358	1 VAL1_BCTV	P1947 abutilon mo
12	216	59.7	358	1 VAL1_BCTV	P14991 beet curly
13	207	57.2	357	1 VAL1_TYLCV	P03567 tomato gold
14	125	34.5	347	1 VAL1_SLCV	P27959 squash leaf
15	64.5	17.8	299	1 Y175_HLELV	Q92mq7 helicohacte
16	62	17.1	285	1 VAL1_TYDAV	P1617 tobacco yell
17	61	16.9	1483	1 CYP1_YEAST	P12351 saccharomyc
18	60.5	16.7	1713	1 LMA3_HUMAN	Q16787 homo sapien
19	50	14.6	127	1 LKSA_DROME	P56079 d phosphati
20	50	14.6	127	1 LKSA_DROME	P56079 d phosphati
21	58.5	16.2	297	1 RHPP_HAHP	P06747 saccharomyc
22	58.5	16.2	299	1 Y175_HLELV	P56112 helicohacte
23	58.5	16.2	367	1 LH44_MOUSE	P33776 mus musculu
24	58.5	16.2	589	1 YE67_HAETN	P45221 haemophilus
25	58	16.0	207	1 IL6_MARMO	O35736 marmota mon
26	58	16.0	256	1 YK09_YEAST	P34247 saccharomyc
27	58	16.0	974	1 RCH_BCOLL	P36682 escherichia
28	58	16.0	974	1 RCH_BCOLL	P36682 escherichia
29	57.5	15.9	136	1 Y452_CAREL	O62350 caenorhadi
30	57	15.7	355	1 CRTP_RHOSH	P54905 rhodobacter
31	56.5	15.6	428	1 GBAL_CANAL	P28868 candida alb
32	56.5	15.6	470	1 RHSA_RHIME	Q923r2 rhizobium m
33	56	15.5	332	1 IPAD_SHIFL	P18013 shigella fl

34 56 15.5 336 1 Y625\_METUA  
35 56 15.5 513 1 HEMO\_CHICK  
36 56 15.5 795 1 LON\_AQUAE  
37 56 15.5 789 1 AFSK\_STRCO  
38 56 15.5 836 1 AFSK\_STRCO  
39 56 15.5 836 1 GSR\_HUMAN  
40 56 15.5 1610 1 CCAD\_MSAU  
41 56 15.5 2161 1 CCAD\_HUMAN  
42 56 15.5 2203 1 CCAD\_RAT  
43 55.5 15.3 201 1 DRPF\_CRAPL  
44 55.5 15.3 224 1 SP2R\_BACSU  
45 55.5 15.3 238 1 RIBE\_ARCFU

## ALIGNMENTS

### RESULT 1

VAL1\_TGMV STANDARD; PRT; 352 AA.  
ID VAL1\_TGMV  
DC P03567  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE AL1 PROTEIN.  
GS ACL.  
QN Tomato golden mosaic virus (TGMV).  
NC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OC  
RP SEQUENCE FROM N.A.  
RA Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.;  
RT "Complete nucleotide sequence of the infectious cloned DNA components  
of tomato golden mosaic virus: potential coding regions and regulatory  
sequences.";  
RT  
RL EMBO J. 3:2197-2205(1984).  
CC  
CC --- SIMILARITY: BELONGS TO: GEMINIVIRUSES AL1 PROTEIN FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed, usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: K02029; -- NOT\_ANNOTATED\_CDS.  
CC PIR: A04170; Q0CVL1.  
CC INTERPRO: IPR001191; --  
CC INTERPRO: IPR001301; --  
CC PFAM: PF00759; Gemin1-ALL; 1.  
CC PRINTS: PR00227; GEMCONTALL.  
CC KX APE-binding; 223 230 ATP (POTENTIAL).  
CC NP\_BIND: 223 230 ATP (POTENTIAL).  
CC SEQUENCE 352 AA; 40332 MW; C33C938E964B4A4 CRC64;

Query Match 98.3%; Score 356; DB 1; Length 352;

Best Local Similarity 98.6%; Pred. No. 1.2e-32;

Matches 69; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TLWGEFQVAGSARGSCOTSDAAALAAALKEALQIREKIPKYLQFPHNLSNL 60

111 TLWGEFQVDRSGKSGCOTSDAAALAAALKEALQIREKIPKYLQFPHNLSNL 170

Db 61 DRIFDKTPPEP 70

171 DRIFDKTPPEP 180

Db 171 DRIFDKTPPEP 180

RESULT 2

VAL1\_PYMV STANDARD; PRT; 361 AA.

ID VAL1\_PYMV

```

AC P27258;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE 01-AUG-1992 (Rel. 23, Last annotation update)
DE 01-AUG-1992 (Rel. 23, Last annotation update)
OS Potato yellow mosaic virus (isolate Venezuela)
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=913111403; PubMed=1856690;
RA Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;
RT "The nucleotide sequence of the infectious cloned DNA components of
RT Potato yellow mosaic virus (P27258)."
RL J. Gen. Virol. 72:1515-1520 (1991).
CC - - - - - SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC - - - - -
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC - - - - -
DE EMBL: D00340; BAA00782.1; -.
DR PIR: J00364; OOCVPT.
DR INTERPRO: IPR001191; -.
DR PRAM: PF00759; GeminA1; 1.
DR PRINTS: PR00227; GEMCONTALL.
DR PRINTS: PR00228; GEMCONTCLV1.
KW ATP-binding.
FT NP-BIND. 222 229 ATP (POTENTIAL).
SQ SEQUENCE 361 AA: 40950 MW: 56273336F1264383 CRC64;

Query Watch 70.7% Score 256; DB 1: Length 361;
Best Local Similarity 69.6%; Pred. No. 1.3e-21;
Matches 48; Conservative 11; Mismatches 10; Indels 0; Gaps 0;

QY 1 TLVGEFOVAGSARGGCGTSDNAAALNASSKEALQIIREKIPKFLFQHNLSNL 60
DB 110 TVMGQFOIDGSRGSGGQTVNANANLNGTGKANKIKERLPERFLFQHNLSNL 169
QY 61 DRIPDKTPE 69
DB 170 DRIPDKAPE 178

RESULT 3
ID VALL_CLVK STANDARD: PRT: 358 AA.
AC P14982;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DE 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN (40.4 KDA PROTEIN).
OS Cassava latent virus (strain West Konyak 844).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP SEQUENCE FROM N.A.
RA Stanley J., Gay M.R.;
RT "Nucleotide sequence of cassava latent virus DNA."
RL Nature 301:260-262 (1983).
CC - - - - - SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC - - - - -
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC or send an email to license@isb-sib.ch).
CC - - - - -
DE EMBL: J02057; -. NOT-ANNOTATED_CDS.
DR INTERPRO: IPR001191; -.
DR PRAM: PF00759; GeminA1; 1.
DR PRINTS: PR00227; GEMCONTALL.
DR PRINTS: PR00228; GEMCONTCLV1.
KW ATP-binding.
FT NP-BIND. 220 227 ATP (POTENTIAL).
SQ SEQUENCE 358 AA: 40346 MW: ED173E753EF92D69 CRC64;

Query Watch 67.1% Score 243; DB 1: Length 358;
Best Local Similarity 62.9%; Pred. No. 5.5e-20;
Matches 44; Conservative 13; Mismatches 13; Indels 0; Gaps 0;

QY 1 TLVGEFOVAGSARGGCGTSDNAAALNASSKEALQIIREKIPKFLFQHNLSNL 60
DB 109 TVMGQFOIDGSRGSGGQTVNANANLNGTGKANKIKERLPERFLFQHNLSNL 169
QY 61 DRIPDKTPE 70
DB 169 DRIPDKPEPP 178

RESULT 4
ID VALL_CLVK CLVW STANDARD: PRT: 358 AA.
AC P14972;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DE 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN (40.4 KDA PROTEIN).
OS Cassava latent virus (strain Nigerian).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP SEQUENCE FROM N.A.
RA Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT "Nucleotide sequence of the infectious cloned DNA components of
RT African cassava mosaic virus (Nigerian strain)."
RL Nucleic Acids Res. 19:1979-1986 (1991).
CC - - - - - SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC - - - - -
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC - - - - -
DE EMBL: X17095; CAA34953.1; -.
DR PIR: S07594; S07594.
DR INTERPRO: IPR001191; -.
DR PRAM: PF00759; GeminA1; 1.
DR PRINTS: PR00227; GEMCONTALL.
DR PRINTS: PR00228; GEMCONTCLV1.
KW ATP-binding.
FT NP-BIND. 220 227 ATP (POTENTIAL).
SQ SEQUENCE 358 AA: 40435 MW: 1D816BB0CB2D5E2C CRC64;

Query Watch 67.1% Score 243; DB 1: Length 358;
Best Local Similarity 62.9%; Pred. No. 5.5e-20;
Matches 44; Conservative 13; Mismatches 13; Indels 0; Gaps 0;

QY 1 TLVGEFOVAGSARGGCGTSDNAAALNASSKEALQIIREKIPKFLFQHNLSNL 60
DB 109 TVMGQFOIDGSRGSGGQTVNANANLNGTGKANKIKERLPERFLFQHNLSNL 169

```

```

OY 61 DRIFDKTPEP 70
DB 169 DRIFQEPAP 178

RESULT 5
VALL TYLVC
ID VALL TYLVC STANDARD; PRT: 362 AA.
AC P36279; 1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE ALL PROTEIN (C1 PROTEIN).
CN C1.
OS Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP MEDLINE=9423446;
RX DRY I.B. Riden, J.E. Drake I.B., Mullineux P.M., Rezaian M.A.;
RA "Nucleotide sequence and genome organization of tomato leaf curl
RT geminivirus.";
RL J. Gen. Virol. 74:147-151(1993).
CC -1: SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: Z25751; CAA81026.1;
DR INTERPRO: IPH001191;
DR INTERPRO: IPH001301;
DR PFAM: PF00799; Geminl_1;
DR PRINTS: PR00227; GEMCOATALL.
DR PRINTS: PR00228; GEMCOATCLVL1.
KW ATP-binding, 221 228 ATP (POTENTIAL).
FT NP_BIND 359 AA; 41065 MW; 2D170A51EF80A3EC CRC64;
SQ SEQUENCE 362 AA: 41197 MW; 343D7184B704098 CRC64;

Query Match 65.28; Score 236; DB 1; Length: 362;
Best Local Similarity 55.38; Pred. No. 3.4e-19;
Matches 47; Conservative 9; Mismatches 13; Indels 16; Gaps 1;

OY 1 TLWGEFQVAGHSARGGQTSNDAAAEALNASSKEALQIREKIPKYLQFHNLNSL 60
DB 111 TLWGEFQIDGRSANGGQSDANDAYAKAINAGSKSEALDVKELAPDYILQFHNLNSL 169
OY 61 DRIFDKTPEP 69
DB 170 DRIFQEPAP 174

RESULT 6
VALL TYLVC
ID VALL TYLVC STANDARD; PRT: 359 AA.
AC P36609;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DE ALL PROTEIN (C1 PROTEIN).
CN C1.
OS Tomato yellow leaf curl virus (strain Murcia) (TYLCV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP MEDLINE=94256836; PubMed=8198442;
RX

RA Noris E., Hidalgo E., Accotto G., Moriones E.;
RT "High similarity among the tomato yellow leaf curl virus isolates
RT from the west Mediterranean basin: the nucleotide sequence of an
RT infectious clone from Spain;
RT Nucleotide sequence of the tomato yellow leaf curl virus isolate
RT from Spain (1994).";
CC -1: SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: Z25751; CAA81026.1;
DR INTERPRO: IPH001191;
DR INTERPRO: IPH001301;
DR PFAM: PF00799; Geminl_1;
DR PRINTS: PR00227; GEMCOATALL.
DR PRINTS: PR00228; GEMCOATCLVL1.
KW ATP-binding, 221 228 ATP (POTENTIAL).
FT NP_BIND 359 AA; 41065 MW; 2D170A51EF80A3EC CRC64;
SQ SEQUENCE 359 AA: 41065 MW; 2D170A51EF80A3EC CRC64;

Query Match 63.38; Score 229; DB 1; Length 359;
Best Local Similarity 60.98; Pred. No. 2e-19;
Matches 42; Conservative 9; Mismatches 18; Indels 0; Gaps 0;

OY 2 LWKGEFQVAGHSARGGQTSNDAAAEALNASSKEALQIREKIPKYLQFHNLNSL 61
DB 111 LWKGEFQIDGRSANGGQSDANDAYAKAINAGSKSEALDVKELAPDYILQFHNLNSL 170
OY 62 DRIFDKTPEP 70
DB 171 DRIFQEPAP 179

RESULT 7
VALL PHIV
ID VALL PHIV STANDARD; PRT: 349 AA.
AC P36609;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE ALL PROTEIN.
CN ALL.
OS Pepper huasteco virus (PHV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RX MEDLINE=94015007; PubMed=8409944;
RA Torres-Pacheco I., Carzon-Tiznado J.A., Herrera-Estrella L.,
RA Rivera-Bustamante R.F.;
RT "Complete nucleotide sequence of pepper huasteco virus: analysis and
RT comparison with bipartite geminiviruses.";
CC -1: SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: X70418; CAA49856.1;
DR PIR: S31875; S31875.
DR PIR: JQ2300; JQ2300.
DR INTERPRO: IPH001191;

```

```

DR INTERPRO: IPR001301;
DR PFAM: PF00799; Gemin1_AL1: 1.
DR PRINTS: PR00227; GEMCOATALL.
KW ATP-binding, 221 228 * ATP (BY SIMILARITY).
FT NP_BIND 221 228 * ATP (BY SIMILARITY).
SQ SEQUENCE 349 AA: 39722 WK; DSFAE76CD56370F4 CRC64;

Query Match
Best local Similarity 62.2%; Score 222; DB 1; Length 349;
Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

QY 1 TLVGEFQVARSARGQQTNDAAALNASSKEALQIREKIPKYLQFPHNLSNL 60
DB 110 TVMEQFQVGRSARGQQTNDAAALNASSKEALQIREKIPKYLQFPHNLSNL 169
QY 61 DRIFOKTPEP 70
DB 170 ERIFKTPPEP 179

RESULT 8
ID VAL1_TYLCM STANDARD: PRT: 359 AA.
AC P27260;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE ALL PROTEIN (C1 PROTEIN).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -|- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
DR EMBL: X61153; CA43466.1;
DR PIR: S22593; S22593.
DR INTERPRO: IPR001191;
DR PFAM: PF00799; Gemin1_AL1: 1.
DR PRINTS: PR00227; GEMCOATALL.
KW ATP-binding, 220 227 * ATP (POTENTIAL).
FT NP_BIND 220 227 * ATP (POTENTIAL).
SQ SEQUENCE 359 AA: 40733 WK; 9717H4A07C38FA7 CRC64;

Query Match
Best local Similarity 61.3%; Score 222; DB 1; Length 359;
Matches 40; Conservative 11; Mismatches 18; Indels 0; Gaps 0;

QY 2 LVNGEFQVARSARGQQTNDAAALNASSKEALQIREKIPKYLQFPHNLSNL 61
DB 111 LVNGEFQVARSARGQQTNDAAALNASSKEALQIREKIPKYLQFPHNLSNL 170
QY 62 DRIFKTPPEP 70
DB 170 ERIFKTPPEP 179

DR INTERPRO: IPR001301;
DR PFAM: PF00799; Gemin1_AL1: 1.
DR PRINTS: PR00227; GEMCOATALL.
KW ATP-binding, 221 228 * ATP (BY SIMILARITY).
FT NP_BIND 221 228 * ATP (BY SIMILARITY).
SQ SEQUENCE 349 AA: 39722 WK; DSFAE76CD56370F4 CRC64;

Query Match
Best local Similarity 60.0%; Score 220; DB 1; Length 353;
Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

QY 1 TLVGEFQVARSARGQQTNDAAALNASSKEALQIREKIPKYLQFPHNLSNL 60
DB 110 TVMEQFQVGRSARGQQTNDAAALNASSKEALQIREKIPKYLQFPHNLSNL 169
QY 61 DRIFOKTPEP 70
DB 170 ERIFKTPPEP 179

RESULT 9
ID VAL1_BGMV STANDARD: PRT: 353 AA.
AC P05175;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE ALL PROTEIN (40.2 KDA PROTEIN).
CC golden mosaic virus
CC Viruses; SSDNA viruses; Geminiviridae; Begomovirus.
RN 11
RA Howarth A.J., Caton J., Bossert M., Goodman R.M.;
RT *Nucleotide sequence of bean golden mosaic virus and a model for gene regulation in geminiviruses.*; 82-3572-3576(1985).
RL Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).
CC -|- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -|- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
DR EMBL: M10070; AAA46318.1;
DR INTERPRO: IPR001191;
DR PFAM: PF00799; Gemin1_AL1: 1.
DR PRINTS: PR00227; GEMCOATALL.
KW ATP-binding, 222 229 * ATP (POTENTIAL).
FT NP_BIND 222 229 * ATP (POTENTIAL).
SQ SEQUENCE 353 AA: 40190 WK; 80FA779DF6029A34 CRC64;

```



-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use, but the copyright and the moral rights are reserved to the EMBL and to the  
 CC Swiss Institute of Bioinformatics. The EMBL and the Swiss Institute of Bioinformatics  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X15983; -. NOT\_ANNOTATED\_CDS.  
 DR PIR; A36214; QCCWML.  
 DR INTERPRO; IPR001191; .  
 DR PFAM; PF00799; Gemini\_L1; 1.  
 DR PRINTS; PR00227; GEMCOATL1.  
 DR PRINTS; PR00228; GEMCOATCL1.  
 KW ATP-Binding. 221 228 ATP (POTENTIAL).  
 FT NP-BIND  
 SQ SEQUENCE 355 AA: 40257 MW: 16426A86A63251E95 CRC64;

Query Match 60.5%; Score 219; DB 1; Length 355;  
 Best Local Similarity 58.6%; Pred. No. 2.6e-17;  
 Matches 41; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

Oy 1 TLWGEFQVARGSGCOTSDAAAEALNASKEEALQIIREKIPKYLEYFQHNLNSNL 60  
 Db 110 TLWGEFQVARGSGCOTSDAAAEALNASKEEALQIIREKIPKYLEYFQHNLNSNL 169  
 Oy 61 DRIFDKTPEP 70  
 Db 170 ERIFAKPEP 179

RESULT 11  
 VAL1\_BCTV STANDARD; PRT; 358 AA.  
 AC P14991;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 01-AUG-1992 (Rel. 23, Last annotation update)  
 DE AL1 PROTEIN (40.8 KDA PROTEIN).  
 CC Best curly top virus (BCTV).  
 CC Viruses; SSDNA viruses; Geminiviridae; Curtovirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Stanley J., Markham P.G., Callis R.J., Pinner M.S.;  
 RT "The nucleotide sequence of an infectious clone of the geminivirus  
 RT BMD J. 5:1761-1767(1986).  
 RL BMD J. 5:1761-1767(1986).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X04144; -. NOT\_ANNOTATED\_CDS.  
 DR INTERPRO; IPR001191; .  
 DR PFAM; PF00799; Gemini\_L1; 1.  
 DR PRINTS; PR00227; GEMCOATL1.  
 DR PRINTS; PR00228; GEMCOATCL1.  
 KW ATP-Binding. 222 229 ATP (POTENTIAL).  
 FT NP-BIND  
 SQ SEQUENCE 358 AA: 40869 MW: 39A45FE3C0B9C333 CRC64;

Query Match 59.7%; Score 216; DB 1; Length 358;  
 Best Local Similarity 55.7%; Pred. No. 5.7e-17;

Matches 39; Conservative 16; Mismatches 15; Indels 0; Gaps 0;  
 Oy 1 TLWGEFQVARGSGCOTSDAAAEALNASKEEALQIIREKIPKYLEYFQHNLNSNL 60  
 Db 110 TLWGEFQVARGSGCOTSDAAAEALNASKEEALQIIREKIPKYLEYFQHNLNSNL 169  
 Oy 61 DRIFDKTPEP 70  
 Db 170 QKIFQPPDP 179

RESULT 12  
 VAL1\_TMOV STANDARD; PRT; 361 AA.  
 AC Q06657;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE AL1 PROTEIN.  
 CC AL1 PROTEIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; L14460; AAC32414.1; .  
 DR PIR; JQ1870; JQ1870.  
 DR INTERPRO; IPR001191; .  
 DR PFAM; PF00799; Gemini\_L1; 1.  
 DR PRINTS; PR00227; GEMCOATL1.  
 DR PRINTS; PR00228; GEMCOATCL1.  
 KW ATP-Binding. 222 229 ATP (BY SIMILARITY).  
 FT NP-BIND  
 SQ SEQUENCE 361 AA: 40516 MW: 8138665CEBAC6950 CRC64;

Query Match 59.7%; Score 216; DB 1; Length 361;  
 Best Local Similarity 55.7%; Pred. No. 5.8e-17;  
 Matches 39; Conservative 16; Mismatches 15; Indels 0; Gaps 0;  
 Oy 1 TLWGEFQVARGSGCOTSDAAAEALNASKEEALQIIREKIPKYLEYFQHNLNSNL 60  
 Db 110 TLWGEFQVARGSGCOTSDAAAEALNASKEEALQIIREKIPKYLEYFQHNLNSNL 169  
 Oy 61 DRIFDKTPEP 70  
 Db 170 ERIFAKPEP 179

RESULT 13  
 VAL1\_TYLCV STANDARD; PRT; 357 AA.  
 AC P27259;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE AL1 PROTEIN (C1 PROTEIN).  
 GN C1.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; L14460; AAC32414.1; .  
 DR PIR; JQ1870; JQ1870.  
 DR INTERPRO; IPR001191; .  
 DR PFAM; PF00799; Gemini\_L1; 1.  
 DR PRINTS; PR00227; GEMCOATL1.  
 DR PRINTS; PR00228; GEMCOATCL1.  
 KW ATP-Binding. 222 229 ATP (BY SIMILARITY).  
 FT NP-BIND  
 SQ SEQUENCE 361 AA: 40516 MW: 8138665CEBAC6950 CRC64;

Query Match 59.7%; Score 216; DB 1; Length 361;  
 Best Local Similarity 55.7%; Pred. No. 5.8e-17;  
 Matches 39; Conservative 16; Mismatches 15; Indels 0; Gaps 0;  
 Oy 1 TLWGEFQVARGSGCOTSDAAAEALNASKEEALQIIREKIPKYLEYFQHNLNSNL 60  
 Db 110 TLWGEFQVARGSGCOTSDAAAEALNASKEEALQIIREKIPKYLEYFQHNLNSNL 169  
 Oy 61 DRIFDKTPEP 70  
 Db 170 ERIFAKPEP 179

RESULT 13  
 VAL1\_TYLCV STANDARD; PRT; 357 AA.  
 AC P27259;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE AL1 PROTEIN (C1 PROTEIN).  
 GN C1.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X04144; -. NOT\_ANNOTATED\_CDS.  
 DR INTERPRO; IPR001191; .  
 DR PFAM; PF00799; Gemini\_L1; 1.  
 DR PRINTS; PR00227; GEMCOATL1.  
 DR PRINTS; PR00228; GEMCOATCL1.  
 KW ATP-Binding. 222 229 ATP (POTENTIAL).  
 FT NP-BIND  
 SQ SEQUENCE 358 AA: 40869 MW: 39A45FE3C0B9C333 CRC64;

Query Match 59.7%; Score 216; DB 1; Length 358;  
 Best Local Similarity 55.7%; Pred. No. 5.7e-17;

```

OS Tomato yellow leaf curl virus (TYLCV).
RN Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RX MEDLINE=91020470; Pubmed=1926771;
RA Navot N., Pichersky E., Zeiden M., Zamir D., Czosnek H.:
RT "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
RL with a single genomic component.";
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the European Bioinformatics
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL: X15556; CA333688.1;
PIR: D40779; QOCVCI.
DR INTERPRO: IPR001191;
DR PFAM: PF00799; Gemini_ALI.1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00228; GEMCOATL1V1.
KW ATP-binding; 319 226 ATP (POTENTIAL).
FT NP_BIND 319 226
SQ SEQUENCE 357 AA; 40678 MW; 539A568E1A63B2A7 CRC64;

Query Match 57.2%; Score 207; DB 1; Length 357;
Best Local Similarity 65.6%; Pred. No. 5.8e-16;
Matches 40; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY 4 WCFVQVAGSRAGCGQTSDNDAEALNASSKEALQIRKIPKYLQFPHNLNSLDRI 63
DB 111 FGVSDIRSGANGQGSANDAYAEALNAGSKSEALNLIKKEAPKDYILOPHNLSSLDRI 170
QY 64 F 64
DB 171 F 171

RESULT 14
VAL1_SUCV STANDARD: PRT; 347 AA.
AC P29048; (REL. 24, Created)
DT 01-DEC-1992 (REL. 24, Last sequence update)
DT 01-DEC-1992 (REL. 24, Last annotation update)
DE AL1 PROTEIN
OS Squash leaf curl virus.
RN Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RX MEDLINE=91082449; Pubmed=1984668;
RA Zeng L., S.G. Lazdins J.B.
RT "Reactivity and localization of the cistidine sequence of the cloned genomic
RT components of a bipartite squash leaf curl geminivirus with a broad
RT host range phenotype.";
RL Virology 180:58-69(1991).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the European Bioinformatics
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL: M36183; AAC32410.1; ALT_INIT.

```

```

DR PIR: C36785; QOCVSI.
DR INTERPRO: IPR001191;
DR PFAM: PF00799; Gemini_ALI.1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00228; GEMCOATL1V1.
KW ATP-binding; 218 225 ATP (POTENTIAL).
FT NP_BIND 218 225
SQ SEQUENCE 347 AA; AFOABDDDE122110F CRC64;

Query Match 34.5%; Score 125; DB 1; Length 347;
Best Local Similarity 37.9%; Pred. No. 1.2e-07;
Matches 25; Conservative 15; Mismatches 22; Indels 4; Gaps 1;

QY 5 GEFQVAGSRAGCGQTSDNDAEALNASSKEALQIRKIPKYLQFPHNLNSLDRI 64
DB 116 GQYKYSQY---GSKSNKDYVHNAVNGAGEALDIKAGDKFTFYVHNLNVERLF 171
QY 65 DKTEPP 70
DB 172 QKPEPP 177

RESULT 15
Y175_HELPJ STANDARD: PRT; 299 AA.
AC Q12901; 2000 (REL. 40, Created)
DT 01-OCT-2000 (REL. 40, Last sequence update)
DT 01-OCT-2000 (REL. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN HP0175 PRECURSOR.
GN HP0175 OR JHP0151.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
RN [1] Helicobacter.
RX MEDLINE=99120557; Pubmed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., Gedonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.F.,
RA Gibson K., Merbery D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
CC -1- SIMILARITY: BELONGS TO THE PPIC/PARVULIN FAMILY OF ROTAMASES.
CC --- STRONG, TO C.JEJUNI CRF2.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the European Bioinformatics
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL: A800145; M305744.1;
PIR: Q13526; LPIN.
DR INTERPRO: IPR00297;
DR PFAM: PF00639; Rotamase; 1.
DR PROSITE: PS01096; PPIC_PPPIASE_1.1.
KW Hypothetical protein; Isomerase; Rotamase; Signal.
FT SIGNAL 1 21
FT CHAIN 152 299
FT REGION 152 299
FT MOTIF 152 299
SQ SEQUENCE 299 AA; 34040 MW; 5C037BCD1110143 CRC64;

Query Match 17.8%; Score 64.5; DB 1; Length 299;
Best Local Similarity 28.1%; Pred. No. 4.2;
Matches 18; Conservative 14; Mismatches 11; Indels 21; Gaps 2;

```

Oy	22	NDAAEALNMS-----SKEEALQITREKIPKYLFOFINLNLSLDR	62
Dd	93	NEAKAEKLNTPEPKAMAEVRKQALVEWAKQAEVVKOTPEKMQDFY--NAKMDQ	150
Oy	63	LFQK 66	
Dd	151	LFVK 154	

Search completed: February 3, 2001, 02:24:29  
Job time: 641 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:17:41 : Search time 118.74 Seconds  
(without alignments)  
40.029 Million cell updates/aec

Title: US-09-289-346a-9

Perfect score: 362

Sequence: 1 TLWGEFQVAGSARGGCGT.....QFHNLSNLDRIFDKTPPEP 70

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 66:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result	No.	Score	Query Match	Length	DB ID	Description
1	356	98.3	352	1	QOCV11	All protein - toma
2	256	70.7	361	1	QOCVPT	All protein - toma
3	243	67.1	358	2	S07594	hypothetical prote
4	236	65.2	362	1	J01887	All protein - toma
5	229	63.3	359	2	S39211	gene C1 protein -
6	225	62.2	349	2	J02300	replicase - pepper
7	225	62.2	359	2	S1875	All protein - pepp
8	225	62.2	359	2	S1875	hypothetical prote
9	219	60.5	351	2	J02327	hypothetical prote
10	219	60.5	355	1	QOCVW1	AV1 protein - abut
11	216	59.7	358	1	J01870	All protein - toma
12	216	59.7	359	2	S39235	gene C1 protein -
13	216	59.7	385	2	S38360	All protein - beet
14	209	57.7	360	2	S59885	All protein - toma
15	209	57.7	360	2	S59885	All protein - toma
16	142	39.2	331	2	S45059	All protein - toma
17	125	34.5	347	1	QOCVSL	All protein - aqua
18	68	18.8	387	2	JC1419	FC gamma (Igc) rec
19	67	18.5	840	2	J36175	probable large ATP
20	64.5	17.8	299	2	B71967	probable peptidyl-
21	62.5	17.3	447	2	T12944	hypothetical prote
22	62	17.1	495	2	T24352	hypothetical prote
23	61.5	17.0	481	2	T24352	probable flagellar
24	61.5	17.0	481	2	T24352	probable flagellar
25	61	16.9	1502	1	RGBNH1	CYC1/CYB3 transcr
26	60.5	16.7	1713	2	A55347	adhesive ligand ep
27	60.5	16.7	2137	2	T05244	hypothetical prote
28	60	16.6	160	2	G82060	hypothetical prote
29	60	16.5	447	2	S52437	CDP-diacylglycerol

probable NADH oxid  
hypothetical prote  
conserved hypothet  
viral mRNA transla  
LAMP1 protein in c  
hypothetical prote  
nonstructural prot  
cell binding facto  
hypothetical prote  
protein kinase bub  
structural mainte  
hypothetical prote  
hypothetical prote  
hypothetical prote  
S-layer protein pr  
hypothetical prote

## ALIGNMENTS

RESULT 1  
QOCV11  
All protein - tomato golden mosaic virus  
C:Species: tomato golden mosaic virus  
A:Note: host Nicotiana sp. (tobacco)  
C:Date: 28-Aug-1985 #sequence\_revision 28-Aug-1985 #text\_change 08-Apr-1994  
E:Accession: A04170 Stein, V.E.; Coutts, R.H.A.; Buck, K.W.  
EMBL J. 3, 2197-2205, 1984  
A:Title: Complete nucleotide sequence of the infectious cloned DNA components of toma  
A:Reference number: A04163  
A:Accession: A04170  
A:Molecule type: DNA  
A:Residues: 1-352 <HAMP>  
C:Comment: The genome consists of two circular, single-stranded DNA components, DNA A  
C:Genetics:  
A:Map position: segment A  
C:Superfamily: Tomato golden mosaic virus All protein

Query Match Best Local Similarity 98.3% Score 356; DB 1; Length 352;  
Matches 69; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 TLWGEFQVAGSARGGCGTSDAAALNASSKEALQIIREKIPKYLQFHNLSN. 60  
Db 111 TLWGEFQVAGSARGGCGTSDAAALNASSKEALQIIREKIPKYLQFHNLSN. 170  
Qy 61 DRIFDKTPPEP 70  
Db 171 DRIFDKTPPEP 180  
RESULT 2  
QOCVPT  
All protein - potato yellow mosaic virus (isolate Venezuela)  
C:Species: potato yellow mosaic virus  
C:Date: Jun-92 #sequence\_revision 30-Jun-1992 #text\_change 16-Jun-2000  
C:Accession: J00364  
E:Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.  
J. Gen. Virol. 72, 1515-1520, 1991  
A:Title: The nucleotide sequence of the infectious cloned DNA components of potato ye  
A:Reference number: J00362; MUID:91311403  
A:Accession: J00364  
A:Status: translation not shown  
A:Residues: 1-351 <QOUP>  
A:Cross-references: GB:D00940; NID:9222458; PIDN:BA00782.1; PID:9222459  
C:Genetics:  
A:Map position: segment A  
C:Superfamily: tomato golden mosaic virus All protein

27.

C>Date: 22-Nov-1993 #sequence\_revision 26-May-1995 #text\_change 20-Sep-1999  
 A:Accession: S31875 Garzon-Tiznado, J.A.; Jimenez, B.; Herrera-Estrella, L.; Rivera-Bu  
 A:Reference number: 531875  
 A:Reference number: 531875  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-349 <TOR>  
 A:Cross-references: EMBL:X70418; NID:g61023; PID:CA49856.1; PID:g61025  
 A:Notes: A new golden mosaic virus  
 C:Superfamily: tomato golden mosaic virus A1 protein

Query Match 62.2%; Score 225; DB 2; Length 349;  
 Best Local Similarity 60.0%; Pred. No. 1e-17;  
 Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

Oy 1 TLVWGEFQVAGSRGCGCOTSDNDAAEALNASKEALQITREKIPKYLQFPHNLNSLD 60  
 Db 111 TLMWGEFQDGRSGAQSQANDATYAKALNSASAEALQIKKEQFQFFLPHNYSNA 169

Oy 61 DRIFDKTPEP 70

Db 170 NRIFQTPPEP 179

# RESULT 8

S22593

hypothetical protein C4 - tomato yellow leaf curl virus

C:Species: tomato yellow leaf curl virus

C>Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 20-Sep-1999

A:Accession: S22593

A:Reference number: 522593

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-359 <RHE>

A:Cross-references: EMBL:X61153; NID:g62211; PID:CA43466.1; PID:g62217

A:Notes: the nucleotide sequence was submitted to the EMBL Data Library, August 1991

C:Superfamily: tomato golden mosaic virus A1 protein

Query Match 61.3%; Score 222; DB 2; Length 359;  
 Best Local Similarity 58.0%; Pred. No. 2.2e-17;  
 Matches 40; Conservative 11; Mismatches 18; Indels 0; Gaps 0;

Oy 2 TLVWGEFQVAGSRGCGCOTSDNDAAEALNASKEALQITREKIPKYLQFPHNLNSLD 61

Db 111 LEMGTFQIDGRSGAQSQANDATYAKALNSAGSQAALDVIKELAPRDLVLFHNINSNLD 170

Oy 62 RIFDKTPEP 70

Db 171 KVFQVPPAP 179

# RESULT 9

JO2327

A1 protein - Indian cassava mosaic virus

N:Alternate names: replication-associated protein

C:Species: Indian cassava mosaic virus

C>Date: 28-Aug-1995 #sequence\_revision 07-Oct-1994 #text\_change 20-Sep-1999

A:Accession: J01870

A:Reference number: 53580

A:Status: Robinson, B.D.

J. Gen. Virol. 74, 2437-2443, 1993

A:Title: Nucleotide sequence evidence for the occurrence of three distinct whitefly-trans

A:Reference number: JO2327

A:Accession: J02327

A:Molecule type: DNA

A:Residues: 1-351 <HON>  
 A:Cross-references: EMBL:224758; NID:g395351; PID:CA480891.1; PID:g584046  
 C:Superfamily: tomato golden mosaic virus A1 protein

Query Match 60.5%; Score 219; DB 2; Length 351;  
 Best Local Similarity 62.7%; Pred. No. 4.8e-17;  
 Matches 42; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

Oy 4 WSEFQVAGSRGCGCOTSDNDAAEALNASKEALQITREKIPKYLQFPHNLNSLDRI 63  
 Db 113 WGTQFQIDGRSGAQSQANDATYAKALNSGSKSEALKTLRELAPROYLQFPHHSSMDRI 172

Oy 64 FDKTPEP 70

Db 173 FTKPPPP 179

# RESULT 10

JO2327

A1 protein - abutilon mosaic virus (isolate West India)

C:Species: abutilon mosaic virus

C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 08-Apr-1994

A:Accession: A36214

A:Reference number: 36214

A:Status: Zimmatt, G.; Jeske, H.

Virol. 178, 461-468, 1990

A:Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as

A:Reference number: A36214; MUID:91020984

A:Accession: A36214

A:Status: A36214

A:Residues: 1-355 <PRI>

A:Cross-references: EMBL:X15983

C:Genetics:

A:Map position: segment A

C:Superfamily: tomato golden mosaic virus A1 protein

Query Match 60.5%; Score 219; DB 1; Length 355;

Best Local Similarity 58.6%; Pred. No. 4.8e-17;

Matches 41; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

Oy 1 TLVWGEFQVAGSRGCGCOTSDNDAAEALNASKEALQITREKIPKYLQFPHNLNSLD 60  
 Db 110 TLMWGEFQIDGRSGAQSQANDATYAKALNSAGSQAALDVIKELAPRDLVLFHNINSNLD 169

Oy 61 DRIFDKTPEP 70

Db 170 ERIFAKAPEP 179

# RESULT 11

JO1870

A1 protein - tomato mottle virus (isolate Florida)

C:Species: tomato mottle virus

C>Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 07-May-1999

A:Accession: J01870

A:Reference number: 1870

A:Status: J. Gen. Virol. 73, 3225-3229, 1992

A:Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated f

A:Reference number: J01869; MUID:93107858

A:Accession: J01870

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-358 <ARO>

A:Cross-references: GB:L14460

C:Genetics:

A:Map position: segment A

C:Superfamily: tomato golden mosaic virus A1 protein

Query Match 59.7%; Score 216; DB 1; Length 358;

Best Local Similarity 55.7%; Pred. No. 1.1e-16;

Matches 39; Conservative 16; Mismatches 15; Indels 0; Gaps 0;

QY 1 TLVWGEFVAGRSARGGCGTSDNAAAEALNASSKEEALQIREKIPKYLFOFHNLSNL 60  
Db 107 TLWGDQIDGRSARGGQGSANDAKANASVQSALAVLREKQDFVLOHNIRSL 166  
QY 61 DRIFDKTPPE 70  
Db 167 ERIFARPEP 176  
RESULT 12  
S59855  
A:Title: CI protein - tomato yellow leaf curl virus  
C:Species: tomato yellow leaf curl virus  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 08-Sep-1997  
R:Accession: S59235  
R:Accession: S59235  
A:Status: preliminary  
A:Residues: 1-359 <R>  
A:Cross-references: EMBL:Z28390; NID:g1041671; PID:g1334964  
C:Superfamily: tomato golden mosaic virus ALL protein

Query Match 59.7%; Score 216; DB 2; Length 359;  
Best Local Similarity 56.5%; Pred. No. 1,le-16;  
Matches 39; Conservative 12; Mismatches 18; Indels 0; Gaps 0;

QY 2 LVWGEFVAGRSARGGCGTSDNAAAEALNASSKEEALQIREKIPKYLFOFHNLSNL 61  
Db 111 LMGFTQIDGRSARGGQGSANDAKANASVQSALAVLREKQDFVLOHNIRSL 170  
QY 62 DRIFDKTPPE 70  
Db 171 KVFQVPPAP 179

RESULT 13  
S59855  
A:Title: CI protein - beet curly top virus  
C:Species: beet curly top virus  
C:Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 20-Sep-1999  
R:Accession: S28360  
R:Accession: S28360  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-385 <STA>  
A:Cross-references: GB:M24597; EMBL:X04144; NID:g210678; PID:AA42751.1; PID:g210679  
C:Superfamily: tomato golden mosaic virus ALL protein

Query Match 59.7%; Score 216; DB 2; Length 385;  
Best Local Similarity 55.7%; Pred. No. 1,le-16;  
Matches 39; Conservative 16; Mismatches 15; Indels 0; Gaps 0;

QY 1 TLVWGEFVAGRSARGGCGTSDNAAAEALNASSKEEALQIREKIPKYLFOFHNLSNL 60  
Db 137 TLWGDQIDGRSARGGQGSANDAKANASVQSALAVLREKQDFVLOHNIRSL 196  
QY 61 DRIFDKTPPE 70  
Db 197 OKIFORPPDP 206

RESULT 14

S59885  
A:Title: replication-associated protein CI - tomato yellow leaf curl virus  
C:Species: tomato yellow leaf curl virus  
C:Date: 14-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 20-Sep-1999  
R:Accession: S59885  
R:Accession: S59885  
A:Status: preliminary  
A:Residues: 1-360 <R>  
A:Cross-references: EMBL:Z40182; NID:g944838; PID:CAA88229.1; PID:g974211  
C:Superfamily: tomato golden mosaic virus ALL protein

Query Match 57.7%; Score 209; DB 2; Length 360;  
Best Local Similarity 60.6%; Pred. No. 6,3e-16;  
Matches 40; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

QY 4 WGEFVAGRSARGGCGTSDNAAAEALNASSKEEALQIREKIPKYLFOFHNLSNIDRI 63  
Db 113 FGVFTQIDGRSARGGQGSANDAKANASVQSALAVLREKQDFVLOHNLSNIDRI 172  
QY 64 FDKTPE 69  
Db 173 FDKTPE 178

RESULT 15

QY 1 AL1 protein - tomato yellow leaf curl virus  
N:Alternate names: CI protein  
C:Species: tomato yellow leaf curl virus  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 23-Jul-1999  
R:Accession: D40779  
R:Accession: D40779  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-357 <NAV>  
A:Cross-references: GB:X15656; NID:g62204; PID:CAA33688.1; PID:g62207  
C:Superfamily: tomato golden mosaic virus ALL protein

Query Match 57.3%; Score 207; DB 1; Length 357;  
Best Local Similarity 65.6%; Pred. No. 1,le-15;  
Matches 40; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY 4 WGEFVAGRSARGGCGTSDNAAAEALNASSKEEALQIREKIPKYLFOFHNLSNIDRI 63  
Db 111 FGVFTQIDGRSARGGQGSANDAKANASVQSALAVLREKQDFVLOHNLSNIDRI 170  
QY 64 F 64  
Db 171 F 171

Search completed: February 3, 2001, 02:17:41  
Job time: 2213 sec





```

Qy 2 LVNFEQVAGRGCGCTSNDAAREALNASSKEEAQIIRKIPKYLFOFHNNNSLD 61
Db 111 LEMFTQDGRSGRGQUTANDAYAKAINAGSKSQALDVIKELAPRDYVLVLFHNNNSLD 170
Qy 62 RIFDKTPEP 70
Db 171 KVFQVPPAP 179

RESULT 2
US-08-809-103B-4
: Sequence 4, Application US/08809103B
: Patent No. 6133505
: GENERAL INFORMATION:
: APPLICANT: GRONENBORN, Bruno
: TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: YOUNG & THOMPSON
: STREET: 745 South 23rd Street
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/809,103B
: FILING DATE: 17-MAR-1997
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: FR 94.11040
: FILING DATE: 15-SEP-1994
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 359 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-809-103B-4

Query Match 61.3%; Score 222; DB 3; Length 359;
Best Local Similarity 58.0%; Pred. No. 7.3e-21;
Matches 40; Conservative 11; Mismatches 18; Indels 0;

Qy 2 LVNFEQVAGRGCGCTSNDAAREALNASSKEEAQIIRKIPKYLFOFHNNNSLD 61
Db 111 LEMFTQDGRSGRGQUTANDAYAKAINAGSKSQALDVIKELAPRDYVLVLFHNNNSLD 170
Qy 62 RIFDKTPEP 70
Db 171 KVFQVPPAP 179

RESULT 3
US-08-809-103B-6
: Sequence 4, Application US/08809103B
: Patent No. 6133505
: GENERAL INFORMATION:
: APPLICANT: GRONENBORN, Bruno
: TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: YOUNG & THOMPSON
: STREET: 745 South 23rd Street
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/809,103B
: FILING DATE: 17-MAR-1997
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: FR 94.11040
: FILING DATE: 15-SEP-1994
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 359 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-809-103B-6

Query Match 61.3%; Score 222; DB 3; Length 359;
Best Local Similarity 58.0%; Pred. No. 7.3e-21;
Matches 40; Conservative 11; Mismatches 18; Indels 0;

Qy 2 LVNFEQVAGRGCGCTSNDAAREALNASSKEEAQIIRKIPKYLFOFHNNNSLD 61
Db 111 LEMFTQDGRSGRGQUTANDAYAKAINAGSKSQALDVIKELAPRDYVLVLFHNNNSLD 170
Qy 62 RIFDKTPEP 70
Db 171 KVFQVPPAP 179

RESULT 4
US-08-809-103B-8
: Sequence 8, Application US/08809103B
: Patent No. 6133505
: GENERAL INFORMATION:
: APPLICANT: GRONENBORN, Bruno
: TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: YOUNG & THOMPSON
: STREET: 745 South 23rd Street
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/809,103B
: FILING DATE: 17-MAR-1997
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: FR 94.11040
: FILING DATE: 15-SEP-1994
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 359 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-809-103B-8

Query Match 61.3%; Score 222; DB 3; Length 359;
Best Local Similarity 58.0%; Pred. No. 7.3e-21;
Matches 40; Conservative 11; Mismatches 18; Indels 0;

Qy 2 LVNFEQVAGRGCGCTSNDAAREALNASSKEEAQIIRKIPKYLFOFHNNNSLD 61
Db 111 LEMFTQDGRSGRGQUTANDAYAKAINAGSKSQALDVIKELAPRDYVLVLFHNNNSLD 170
Qy 62 RIFDKTPEP 70
Db 171 KVFQVPPAP 179

```

```

: Sequence 6, Application US/08809103B
: Patent No. 6133505
: GENERAL INFORMATION:
: APPLICANT: GRONENBORN, Bruno
: TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: YOUNG & THOMPSON
: STREET: 745 South 23rd Street
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/809,103B
: FILING DATE: 17-MAR-1997
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: FR 94.11040
: FILING DATE: 15-SEP-1994
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 359 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-809-103B-6

Query Match 61.3%; Score 222; DB 3; Length 359;
Best Local Similarity 58.0%; Pred. No. 7.3e-21;
Matches 40; Conservative 11; Mismatches 18; Indels 0;

Qy 2 LVNFEQVAGRGCGCTSNDAAREALNASSKEEAQIIRKIPKYLFOFHNNNSLD 61
Db 111 LEMFTQDGRSGRGQUTANDAYAKAINAGSKSQALDVIKELAPRDYVLVLFHNNNSLD 170
Qy 62 RIFDKTPEP 70
Db 171 KVFQVPPAP 179

RESULT 4
US-08-809-103B-8
: Sequence 8, Application US/08809103B
: Patent No. 6133505
: GENERAL INFORMATION:
: APPLICANT: GRONENBORN, Bruno
: TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: YOUNG & THOMPSON
: STREET: 745 South 23rd Street
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/809,103B
: FILING DATE: 17-MAR-1997
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: FR 94.11040
: FILING DATE: 15-SEP-1994
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 359 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-809-103B-8

Query Match 61.3%; Score 222; DB 3; Length 359;
Best Local Similarity 58.0%; Pred. No. 7.3e-21;
Matches 40; Conservative 11; Mismatches 18; Indels 0;

Qy 2 LVNFEQVAGRGCGCTSNDAAREALNASSKEEAQIIRKIPKYLFOFHNNNSLD 61
Db 111 LEMFTQDGRSGRGQUTANDAYAKAINAGSKSQALDVIKELAPRDYVLVLFHNNNSLD 170
Qy 62 RIFDKTPEP 70
Db 171 KVFQVPPAP 179

```

```

1 COUNTRY: U.S.A.
2
3 COMPUTER READABLE FORM:
4 MEDIUM TYPE: Floppy disk
5 OPERATING SYSTEM: PC-DOS/MS-DOS
6 CURRENT APPLICATION DATA:
7   APPLICATION NUMBER: US/08/809,103B
8   FILING DATE: 17-MAR-1997
9 CLASSIFICATION: 800
10 PRIOR APPLICATION DATA:
11   APPLICATION NUMBER: FR 94.11040
12   FILING DATE: 15-SEP-1994
13 PRIOR APPLICATION DATA:
14   APPLICATION NUMBER: WO PCT/FR95/01102
15   FILING DATE: 15-SEP-1995
16 ATTORNEY/AGENT INFORMATION:
17   NAME: PATCH, Andrew J, 2, 925
18   REFERENCE/DOCKET NUMBER: US94AL CNR TOM
19 TELECOMMUNICATION INFORMATION:
20   TELEPHONE: (703) 521-2297
21   TELEX: 248425 EMBON
22 INFORMATION FOR SEQ ID NO: 8:
23   SEQUENCE CHARACTERISTICS:
24     LENGTH: 359 amino acids
25     TYPE: amino acid
26     TOPOLOGY: linear
27 MOLECULE TYPE: protein
28 US-08-809-103B-8
29
30 Query Match          61.3%; Score 222; DB 3; Length 359;
31 Best Local Similarity 58.0%; Pred. No. 7,38-21;
32 Mismatch 40; Conservative 11; Mismatches 18; Indels 0; Gaps 0;
33
34 Oy 2 LWMGEFQVAGRSARGCQTSNDAAEALNAGSKSALQIPEKYLFOFHNLSNLD 61
35   |||||:|||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
36 Db 111 LEMGTQIDGRSARGCQOTANDAYAKATNAGSKSALQIVIKELAPRDYVYLFHFNLSNLD 170
37
38 Oy 62 RFDKPTPEP 70
39
40 Db 171 KVFQVPPAP 179
41
42 RESULT 5
43 US-08-600-982-24
44 Sequence 24, Application US/08/600982
45 Patent No. 61209910N
46 GENERAL INFORMATION:
47   APPLICANT: Carter, William G.
48   APPLICANT: GIL, Susanna A.
49   APPLICANT: Ryan, Maureen C.
50   TITLE OF INVENTION: Epiligrin, an Epithelial Ligand for
51   NUMBER OF SEQUENCES: 30
52   CORRESPONDENCE ADDRESS:
53     ADDRESSEE: Christensen, O'Connor, Johnson, and Kindness
54     STREET: 1420 Fifth Avenue
55     CITY: Seattle
56     STATE: WA
57     COUNTRY: USA
58     ZIP: 98101-8100
59 COMPUTER READABLE FORM:
60 MEDIUM TYPE: Floppy disk
61 COMPUTER: IBM PC compatible
62 OPERATING SYSTEM: PC-DOS/MS-DOS
63 SOFTWARE: Patent Release #1.0, Version #1.25
64 CURRENT APPLICATION DATA:
65   APPLICATION NUMBER: US/08/600,982
66   FILING DATE: 02-SEP-1994
67 CLASSIFICATION:
68 ATTORNEY/AGENT INFORMATION:
69   NAME: Shelton, Dennis K.
70   REGISTRATION NUMBER: 26,997
71 TELECOMMUNICATION INFORMATION:
72   TELEPHONE: (206) 682-8100
73   TELEFAX: (206) 682-0779
74 INFORMATION FOR SEQ ID NO: 24:
75   SEQUENCE CHARACTERISTICS:
76     LENGTH: 1713 amino acids
77     TYPE: amino acid
78     TOPOLOGY: linear
79 MOLECULE TYPE: protein
80 DESCRIPTION: E170 protein as translated from sequence of
81   DESCRIPTION: FIGURES 15A-15F, and as shown also in FIGURES 19A-19R
82 PCT-US94-10261A-24

```

```

1 CLASSIFICATION: 435
2 ATTORNEY/AGENT INFORMATION:
3   NAME: Shelton, Dennis K.
4   REGISTRATION NUMBER: 26,997
5 TELECOMMUNICATION INFORMATION:
6   TELEPHONE: (206) 682-8100
7   TELEFAX: (206) 224-0779
8 INFORMATION FOR SEQ ID NO: 24:
9   SEQUENCE CHARACTERISTICS:
10     LENGTH: 1713 amino acids
11     TYPE: amino acid
12     TOPOLOGY: linear
13 MOLECULE TYPE: protein
14 DESCRIPTION: E170 protein as translated from sequence
15   DESCRIPTION: of FIGURES 15A-15F, and as shown also in FIGURES
16   DESCRIPTION: 19A-19R
17 US-08-600-982-24
18
19 Query Match          16.7%; Score 60.5; DB 3; Length 1713;
20 Best Local Similarity 39.1%; Pred. No. 40;
21 Mismatch 18; Conservative 10; Mismatches 11; Indels 7; Gaps 3;
22
23 Oy 24 AAALN--ASSKEALQ--LLREKIPEKYLFOFHNLSNLDRIEDK 66
24   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
25 Db 568 AAEDAANAASASALQIVIKEDLPK---AKTLSSNSDKLLNE 609
26
27 RESULT 6
28 PCT-US94-10261A-24
29 Sequence 24, Application PCT/US9410261A
30 GENERAL INFORMATION:
31   APPLICANT: Carter, William G.
32   APPLICANT: GIL, Susanna A.
33   APPLICANT: Ryan, Maureen C.
34   TITLE OF INVENTION: Epiligrin, an Epithelial Ligand for
35   NUMBER OF SEQUENCES: 30
36   CORRESPONDENCE ADDRESS:
37     ADDRESSEE: Christensen, O'Connor, Johnson, and Kindness
38     STREET: 1420 Fifth Avenue
39     CITY: Seattle
40     STATE: WA
41     COUNTRY: USA
42     ZIP: 98101-8100
43 COMPUTER READABLE FORM:
44 MEDIUM TYPE: Floppy disk
45 COMPUTER: IBM PC compatible
46 OPERATING SYSTEM: PC-DOS/MS-DOS
47 SOFTWARE: Patent Release #1.0, Version #1.25
48 CURRENT APPLICATION DATA:
49   APPLICATION NUMBER: PCT/US94/10261A
50   FILING DATE: 02-SEP-1994
51 CLASSIFICATION:
52 ATTORNEY/AGENT INFORMATION:
53   NAME: Shelton, Dennis K.
54   REGISTRATION NUMBER: 26,997
55 TELECOMMUNICATION INFORMATION:
56   TELEPHONE: (206) 682-8100
57   TELEFAX: (206) 224-0779
58 INFORMATION FOR SEQ ID NO: 24:
59   SEQUENCE CHARACTERISTICS:
60     LENGTH: 1713 amino acids
61     TYPE: amino acid
62     TOPOLOGY: linear
63 MOLECULE TYPE: protein
64 DESCRIPTION: E170 protein as translated from sequence of
65   DESCRIPTION: FIGURES 15A-15F, and as shown also in FIGURES 19A-19R
66 PCT-US94-10261A-24

```

```

Query Match          16.7%; Score 60.5; DB 4; Length 1713;
Best Local Similarity 39.1%; Pred. No. 40;

```



```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT PATENTIN RELEASE #1.0, Version #1.25
; APPLICATION NUMBER: US/08/446.100
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hendricks, Glenna
; REGISTRATION NUMBER: 32,535
; TELEPHONE: (703) 425-2767
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 425-4250
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 454 amino acids
; TYPE: amino acid
; STRIDE: 1 residue/1 single
; TOPOLOGY: single
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ANTI-SENSE: YES
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: human esterases
; US-08-446-100-31

Query Match 15.5%: Score 56; DB 3; Length 454;
Best Local Similarity 32.3%: Pred. No. 28;
Matches 20; Conservative 10; Mismatches 22; Indels 10; Gaps 4;

Qy 17 GQTSNDAA-AEALNASSKEAQLIIRKIPKYLQPF-----HNLSNMLD-RFDXT 67
Db 160 GCKTTTSAWVCLRQKTEELLETTL-KIGNSLKYRTQRESHLLGYVILDMILLAT 218

Qy 68 PE 69
Db 219 PE 220

RESULT 10
US-08-419-652-6
; Sequence 6, Application US/08419652
; Patent No. 5831007
; GENERAL INFORMATION:
; APPLICANT: Chua, Anne O
; APPLICANT: Gubler, Ulrich A
; TITLE OF INVENTION: INTERLEUKIN-12 RECEPTOR
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: United States of America
; COMPUTER: 3110-119 FORM;
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/419,652
; FILING DATE: 11-APR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: US 08/248,532
; FILING DATE: 31-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/094,713

```

```

; FILING DATE: 19-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Kass, Alan P
; REGISTRATION NUMBER: 32142
; TELEPHONE: (201) 235-4205
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-4205
; TELEFAX: (201) 235-3500
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 602 amino acids
; TYPE: amino acid
; STRIDE: 1 residue/1 single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; NAME/KEY: Region
; LOCATION: 1..602
; OTHER INFORMATION: /note= *Represents residues 98 to
; OTHER INFORMATION: 731 of human granulocyte colony-stimulating
; factor-receptor.*
; US-08-419-652-6

Query Match 15.5%: Score 56; DB 2; Length 602;
Best Local Similarity 24.0%: Pred. No. 40;
Matches 12; Conservative 9; Mismatches 23; Indels 6; Gaps 1;

Qy 3 FOVAGRSARGCQTSNDAAAEALNASSKEAQLIIRKIPKYLQPHNL 56
Db 37 FFLASFKSGCNQTCGSLDLYPKDQSHCC-----LPRALLLYQNM 80

RESULT 11
US-07-923-976-6
; Sequence 6, Application US/07923976
; Patent No. 5574136
; GENERAL INFORMATION:
; APPLICANT: Nagata, Shigeazu
; APPLICANT: Fukunaga, Rikio
; TITLE OF INVENTION: DNA Encoding Granulocyte
; TITLE OF INVENTION: Colony-Stimulating Factor Receptor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones, Tuller & Cooper, P.C.
; STREET: One Box 2266 Eads Station
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/923,976
; FILING DATE: 19920922
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 74539/1990
; FILING DATE: 23-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 176629/1990
; FILING DATE: 03-JUL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP91/00375
; FILING DATE: 22-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hendricks, Glenna W.
; REGISTRATION NUMBER: 32,535
; TELEPHONE: (703) 425-2767
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-415-1500
; TELEFAX: 703-415-1508

```

INFORMATION FOR SEQ ID NO: 6:  
LENGTH: 771 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-923-976-6

Query Match 15.5%: Score 56; DB 1; Length 771;  
Best Local Similarity 24.0%: Pred. No. 55;  
Matches 12; Conservative 9; Mismatches 23; Indels 6; Gaps 1;  
QY 7 FOVAGSARGCQTSNDAAAEALNASKEEALQIREKIPEKYLQFPHNL 56  
DB 156 FTLSKFSKSGNCQTQGSILDCVPGQGSGCC-----IPRKHLLLYQNM 199

RESULT 12  
542248-2  
PATENT NO. 542248  
APPLICANT: SMITH, CRAIG A.; LARSEN, ALF D.; SIMS, JOHN E.;  
BENSON, CURTIS M.  
TITLE OF INVENTION: DNA SEQUENCES ENCODING GRANULOCYTE-COLONY  
STIMULATING FACTOR RECEPTORS  
NUMBER OF SEQUENCES: 6  
CURRENT APPLICATION DATA:  
PRIORITY NUMBER: 587,329  
FILING DATE: 15-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 587,329  
FILING DATE: 24-SEP-1990  
APPLICATION NUMBER: 522,952  
FILING DATE: 03-APR-1990  
APPLICATION NUMBER: 416,306  
FILING DATE: 26-SEP-1989  
APPLICATION NUMBER: 413,816  
FILING DATE: 26-SEP-1989  
SEQ ID NO: 2;  
LENGTH: 783  
542248-2

Query Match 15.5%: Score 56; DB 5; Length 783;  
Best Local Similarity 24.0%: Pred. No. 56;  
Matches 12; Conservative 9; Mismatches 23; Indels 6; Gaps 1;  
QY 7 FOVAGSARGCQTSNDAAAEALNASKEEALQIREKIPEKYLQFPHNL 56  
DB 156 FTLSKFSKSGNCQTQGSILDCVPGQGSGCC-----IPRKHLLLYQNM 199

RESULT 13  
US-07-923-976-4  
Sequence 4, Application US/07923976  
Patent No. 5574136  
GENERAL INFORMATION:  
APPLICANT: Negata, Shigekazu  
ATTORNEY/AGENT INFORMATION:  
TITLE OF INVENTION: DNA Encoding Granulocyte  
Stimulating Factor Receptor  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Jones, Tullar & Cooper, P.C.  
STREET: P.O. Box 2266 Eads Station  
CITY: Arlington  
STATE: Virginia  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CUMULATIVE SEQUENCE ALIGNMENT DATA: JP 176629/1990  
APPLICATION NUMBER: JP 176629/1990  
FILING DATE: 03-JUL-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP91/00375  
FILING DATE: 22-MAR-1991  
ATTORNEY/AGENT INFORMATION:

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/923-976  
FILING DATE: 19920922  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 74539/1990  
FILING DATE: 23-MAR-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 176629/1990  
FILING DATE: 03-JUL-1990  
APPLICATION NUMBER: PCT/JP91/00375  
FILING DATE: 22-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Bellwege, James W.  
REGISTRATION NUMBER: 28 808  
REFERENCE/DOCKET NUMBER: 514853  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-415-1500  
TELEFAX: 703-415-1508  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 836 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-923-976-4

Query Match 15.5%: Score 56; DB 1; Length 836;  
Best Local Similarity 24.0%: Pred. No. 61;  
Matches 12; Conservative 9; Mismatches 23; Indels 6; Gaps 1;  
QY 7 FOVAGSARGCQTSNDAAAEALNASKEEALQIREKIPEKYLQFPHNL 56  
DB 156 FTLSKFSKSGNCQTQGSILDCVPGQGSGCC-----IPRKHLLLYQNM 199

RESULT 14  
US-07-923-976-8  
Sequence 8, Application US/07923976  
Patent No. 5574136  
GENERAL INFORMATION:  
APPLICANT: Negata, Shigekazu  
ATTORNEY/AGENT INFORMATION:  
TITLE OF INVENTION: DNA Encoding Granulocyte  
Stimulating Factor Receptor  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Jones, Tullar & Cooper, P.C.  
STREET: P.O. Box 2266 Eads Station  
CITY: Arlington  
STATE: Virginia  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CUMULATIVE SEQUENCE ALIGNMENT DATA: JP 176629/1990  
APPLICATION NUMBER: JP 176629/1990  
FILING DATE: 03-JUL-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP91/00375  
FILING DATE: 22-MAR-1991  
ATTORNEY/AGENT INFORMATION:

NAME: Hellwege, James W.  
REGISTRATION NUMBER: 28,808  
REFERENCE/DOCKET NUMBER: 514853  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-415-1500  
TELEFAX: 703-415-1508  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 863 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-323-976-8

Query Match 15.5% Score 56; DB 1; Length 863;  
Best Local Similarity 24.0%; Pred. No. 64;  
Matches 12; Conservative 9; Mismatches 23; Indels 6; Gaps 1;

Qy 7 FOVAGRSANGQSTSDNMAEALNASSKEALQIIREKIPKYLQPHNL 56  
Db 156 FTLASFSGNQQTQGDSTILQVPGDSHCC-----IPRKHLLYQNM 199

RESULT 15  
US-07-745-206A-2  
; Sequences 2, Application US/07745206A  
; Patent No. 5299211  
; GRANTING JURISDICTION:  
; APPLICANT: Harpold, Michael  
; APPLICANT: Ellis, Steven  
; APPLICANT: Williams, Mark  
; APPLICANT: McCue, Ann  
; APPLICANT: Feldman, Daniel  
; TITLE OF INVENTION: Human Calcium Channel Compositions and  
; METHODS OF INVENTION: Methods  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fitch, Even, Tabin & Flannery  
; STREET: 135 S. LaSalle  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: U.S.A.  
; ZIP CODE: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/745,206A  
; PRIORITY: 07/191615  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Feder, Scott B  
; REFERENCE/DOCKET NUMBER: 51504  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-372-7842  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2161 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-745-206A-2

Query Match 15.5% Score 56; DB 1; Length 2161;  
Best Local Similarity 51.9%; Pred. No. 211;02;  
Matches 14; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 22 NDAAEALNASSKEALQIIREKIPK 48  
| | | | | : | | | | : | | | |

Db 754 NLDAESLNTAQKDEAEKERKRIARK 780  
Search completed: February 3, 2001, 02:19:51  
Job time: 1719 sec

---

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
 Copyright (c) 1993 - 2000 CompuGen Ltd.  
 OW protein - protein search, using sw model  
 Run on: February 3, 2001, 02:15:29 ; Search time 144.12 seconds  
 (without alignments)  
 16,608 Million cell updates/sec

Title: US-09-289-346a-9  
 Perfect score: 362  
 Sequence: 1 TLVNGFQVAGRSARGQQT.....TQFHNLNSLDRFDKTEP 70

Scoring table:  
 BLOSUM62  
 Gapop 10.0, Gapext 0.5

Searched: 268485 seqs, 34193795 residues  
 Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: A\_Geneseq\_36.\*

1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
 2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
 3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
 4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
 5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
 6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
 7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
 8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
 9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
 10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
 11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
 12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
 13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
 14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
 15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
 16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
 17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
 18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
 19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
 20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
 21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query		ID	Description
		Match	Length		
1	222	61.3	353	18 W34338	Bean golden mosaic
2	222	61.3	353	18 W34332	Bean golden mosaic
3	222	61.3	353	18 W34333	Bean golden mosaic
4	222	61.3	353	18 W34334	Bean golden mosaic
5	222	61.3	353	18 W34335	Bean golden mosaic
6	222	61.3	353	17 W34336	Sardinian tomato Y
7	222	61.3	359	17 W88871	Sardinian tomato Y
8	222	61.3	359	17 W88872	Sardinian tomato Y
9	220	60.8	353	8 P70407	OHF 4 gene product
10	220	60.8	361	18 W34336	Tomato mottle virus
11	220	60.8	361	18 W34324	Tomato mottle virus
12	220	60.8	361	18 W34325	Tomato mottle virus

## SUMMARIES

## ALIGNMENTS

RESULT	1	220	60.8	361	18	W34326
W34338	14	213.5	59.0	361	8	P70562
ID	15	208	57.5	362	19	W56495
XX	16	207	57.2	357	18	W34329
XX	17	207	57.2	357	18	W34330
XX	18	207	57.2	357	18	W34331
XX	19	199	55.0	357	18	W34337
XX	20	65	18.0	335	21	Y92317
XX	21	65	18.0	335	21	Y92318
XX	22	64.5	17.8	299	18	W55450
XX	23	64.5	17.8	512	19	W68473
XX	24	62.5	17.3	447	21	Y89889
XX	25	62.5	17.3	447	21	Y89890
XX	26	62.5	17.3	447	21	Y50947
XX	27	62.5	17.3	456	21	Y87190
XX	28	60.5	16.7	204	20	W89814
XX	29	60.5	16.7	299	20	W89829
XX	30	60.5	16.7	299	20	W89849
XX	31	60.5	16.7	1713	16	R70148
XX	32	58.5	16.2	189	21	Y95043
XX	33	58.5	16.2	299	19	W93322
XX	34	58.5	16.2	299	19	W93322
XX	35	58.5	16.2	299	19	W73034
XX	36	58	16.0	844	18	W05773
XX	37	57	15.7	108	16	R82847
XX	38	56.5	15.6	249	20	W89822
XX	39	56.5	15.6	269	20	W89860
XX	40	56.5	15.6	304	19	W98662
XX	41	56	15.5	206	20	Y32577
XX	42	56	15.5	206	20	Y32577
XX	43	56	15.5	213	17	W00648
XX	44	56	15.5	333	17	R99141
XX	45	56	15.5	454	21	Y49470

RESULT 1

ID W34338 standard; Protein: 353 AA.

XX W34338;

XX W34338;

XX 27-APR-1998 (first entry)

XX Bean golden mosaic geminivirus C1 protein.

XX Geminivirus; BMV; C1 gene; transdominant mutation;

XX Transgenic plant; disease resistance.

XX Bean golden mosaic virus type II isolate Guatemala.

XX W057739110-AL.

XX 23-OCT-1997.

XX 15-APR-1997; 97WO-US050300.

XX 16-APR-1996; 96US-0015517.

XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.

XX (WISC ) WISCONSIN AIDMNI RES FOUND.

XX Ahlquist PG, Hanson SF, Lou HT, Maxwell DP, Stout JT;

XX WPI: 1997-526447/48.

XX N F505; Y93514.

XX Transgenic plants expressing geminivirus AC1 and C1 wild-type and mutant genes - have increased resistance to geminivirus infection e.g. tomato mottle virus, tomato yellow leaf curl virus or bean golden mosaic geminivirus

XX Example 5; Page 100-102; 132pp; English.

XX This sequence comprises the bean golden mosaic virus (BGMV) CI

XX production of transgenic plants containing DNA comprising CI or AC1

XX wild-type or mutant sequences that negatively interfere in trans

XX with geminiviral replication during infection. Such transgenic

XX plants are resistant to viral infection. The AC1/CI genes are

XX especially from BGMV, tomato mottle virus or tomato yellow leaf

XX curl virus (see #93282-93) and encode polypeptides (see #34324-35)

XX that have mutations in the highly conserved DNA-nicking domain

XX and/or the NTP-binding domains.

XX Sequence 353 AA;

SQ

Query Match 61.3%; Score 222; DB 18; Length 353;

Best Local Similarity 60.0%; Pred. No. 7, 2e-20;

Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

OY 1 TLVKEFGVAGRSARGGQTSNDAAAEALNASSKEALQIIREKIPKYLFOFINLSNL 60

DB 110 TLVKEFGVAGRSARGGQTSNDAAAEALNASSKEALQIIREKIPKYLFOFINLSNL 60

OY 61 DRIFDKTPEP 70

DB 170 erifvkvpep 179

RESULT 2

W34332

ID W34332 standard; Protein; 353 AA.

XX W34332;

XX W34332;

DT 27-APR-1998 (first entry)

DE Bean golden mosaic geminivirus C1 BGA190 mutant.

XX Geminivirus; BGMV; CI gene; transdominant mutation;

XX transgenic plant; disease resistance.

OS Bean golden mosaic virus type II isolate Guatemala.

XX W09739110-A1.

PD 23-OCT-1997.

XX 15-APR-1997; 97WQ-US06300.

PR 16-APR-1996; 96US-0015517.

PA (SEMI-) SEMINIS VEGETABLE SEEDS INC.

XX Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;

PI WPI; 1997-526447/48.

XX N-PSDB; T93290.

XX Transgenic plants expressing geminivirus AC1 and CI wild-type and

XX mutant genes - have increased resistance to geminivirus infection

XX e.g. tomato mottle virus, tomato yellow leaf curl virus or bean

XX golden mosaic geminivirus

XX Example 5; Page 103-105; 132pp; English.

XX This protein comprises a control mutant of the bean golden mosaic

XX virus (BGMV) CI protein (see #34338) that is required for

XX replication. It is encoded by mutated CI open reading frame

XX #BAC190 (see #93290). The invention involves production of

XX transgenic plants containing DNA comprising geminivirus CI or AC1

CC wild-type or mutant sequences that negatively interfere in trans

CC with geminiviral replication during infection. Such transgenic

CC plants are resistant to viral infection. The AC1/CI genes are

CC especially from BGMV, tomato mottle virus or tomato yellow leaf

CC curl virus (see #93282-93) and encode polypeptides (see #34324-35)

CC that have mutations in the highly conserved DNA-nicking domain

CC and/or the NTP-binding domains.

XX Sequence 353 AA;

SQ

Query Match 61.3%; Score 222; DB 18; Length 353;

Best Local Similarity 60.0%; Pred. No. 7, 2e-20;

Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

OY 1 TLVKEFGVAGRSARGGQTSNDAAAEALNASSKEALQIIREKIPKYLFOFINLSNL 60

DB 110 TLVKEFGVAGRSARGGQTSNDAAAEALNASSKEALQIIREKIPKYLFOFINLSNL 60

OY 61 DRIFDKTPEP 70

DB 170 erifvkvpep 179

RESULT 3

W34333

ID W34333 standard; Protein; 353 AA.

XX W34333;

XX W34333;

DT 27-APR-1998 (first entry)

DE Bean golden mosaic geminivirus C1 BGA221 mutant.

XX Geminivirus; BGMV; CI gene; transdominant mutation;

XX transgenic plant; disease resistance.

OS Bean golden mosaic virus type II isolate Guatemala.

XX W09739110-A1.

PD 23-OCT-1997.

XX 15-APR-1997; 97WQ-US06300.

PR 16-APR-1996; 96US-0015517.

PA (SEMI-) SEMINIS VEGETABLE SEEDS INC.

XX Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;

PI WPI; 1997-526447/48.

XX N-PSDB; T93291.

XX Transgenic plants expressing geminivirus AC1 and CI wild-type and

XX mutant genes - have increased resistance to geminivirus infection

XX e.g. tomato mottle virus, tomato yellow leaf curl virus or bean

XX golden mosaic geminivirus

XX Example 5; Page 107-109; 132pp; English.

XX This protein comprises a transdominant lethal mutant of the bean

XX golden mosaic virus (BGMV) CI protein (see #34338) that is required

XX for replication. It is encoded by mutated CI open reading frame

XX #BAC221 (see #93291) and carries a mutation in the NTP-binding

XX domain of the protein. The invention involves production of transgenic

XX plants containing DNA comprising geminivirus CI or AC1 wild-type or mutant

XX sequences that negatively interfere in trans with geminiviral

XX replication during infection. Such transgenic plants are resistant

XX to viral infection. The AC1/CI genes are especially from BGMV,

XX tomato mottle virus or tomato yellow leaf curl virus (see #93282-93)

XX and encode polypeptides (see #34324-35) that have mutations in the

CC highly conserved DNA-nicking and/or the NTP-binding domains.

XX Sequence 353 AA;

Query Match 61.3%; Score 222; DB 18; Length 353;

Best Local Similarity 60.0%; Pred. No. 7.2e-20;

Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

Oy 1 TLWGEFVQAGRSAGCCTSDAAAEALNASSKEALQIIRKIPKYLQFHNLSNL 60

Db 110 TLWGEFVQAGRSAGCCTSDAAAEALNASSKEALQIIRKIPKYLQFHNLSNL 169

Oy 61 DRIFDXTPEP 70

Db 170 erifvkvpep 179

RESULT 4

ID W34334 standard; Protein; 353 AA.

AC W34334;

DT 27-APR-1998 (first entry)

DE Bean golden mosaic geminivirus Cl PGA228 mutant.

KW Geminivirus; BGWV; Cl gene; transdominant mutation;

KW transgenic plant; disease resistance.

OS Bean golden mosaic virus type II isolate Guatemala.

XX W09739110-Al.

XX 23-OCT-1997.

PF 15-APR-1997; 97WO-US06300.

PR 16-APR-1996; 96US-0015517.

XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.

PA (WISC) WISCONSIN ALMONI RES FOUND.

XX Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;

XX WPI: 1997-526447/48.

XX N-PSDB: T93292.

XX Transgenic plants expressing geminivirus AC1 and Cl wild-type and

mutant genes have increased resistance to geminivirus infection

PT e-9 tomato mottle virus and tomato yellow leaf curl virus or bean

PT golden mosaic geminivirus

XX Example 5; Page 111-112; 132pp; English.

XX This protein comprises a transdominant lethal mutant of the Bean

golden mosaic virus (BGWV) Cl protein (see W34338) that is required

for replication. It is encoded by mutated Cl open reading frame

CC BGAC262 (see T93293) and carries a mutation in the NTP-binding

CC domain. The invention involves production of transgenic plants

CC containing DNA comprising geminivirus Cl or AC1 wild-type or mutant

CC sequences that negatively interfere in trans with geminiviral

CC replication during infection. Such transgenic plants are resistant

CC to viral infection. The AC1/Cl genes are especially from BGWV.

CC tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)

CC and encode polypeptides (see W34324-35) that have mutations in the

CC highly conserved DNA nicking and/or the NTP-binding domains.

XX Sequence 353 AA;

Query Match 61.3%; Score 222; DB 18; Length 353;

Best Local Similarity 60.0%; Pred. No. 7.2e-20;

Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

Oy 1 TLWGEFVQAGRSAGCCTSDAAAEALNASSKEALQIIRKIPKYLQFHNLSNL 60

Db 110 TLWGEFVQAGRSAGCCTSDAAAEALNASSKEALQIIRKIPKYLQFHNLSNL 169

Oy 61 DRIFDXTPEP 70

Db 170 erifvkvpep 179

RESULT 5

ID W34335 standard; Protein; 353 AA.

AC W34335;

DT 27-APR-1998 (first entry)

DE Bean golden mosaic geminivirus Cl BGA262 mutant.

KW Geminivirus; BGWV; Cl gene; transdominant mutation;

KW transgenic plant; disease resistance.

Best Local Similarity 60.0%; Pred. No. 7.2e-20;

Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

Oy 1 TLWGEFVQAGRSAGCCTSDAAAEALNASSKEALQIIRKIPKYLQFHNLSNL 60

Db 110 TLWGEFVQAGRSAGCCTSDAAAEALNASSKEALQIIRKIPKYLQFHNLSNL 169

Oy 61 DRIFDXTPEP 70

Db 170 erifvkvpep 179

RESULT 5

ID W34335 standard; Protein; 353 AA.

AC W34335;

DT 27-APR-1998 (first entry)

DE Bean golden mosaic geminivirus Cl BGA262 mutant.

KW Geminivirus; BGWV; Cl gene; transdominant mutation;

KW transgenic plant; disease resistance.

OS Bean golden mosaic virus type II isolate Guatemala.

XX W09739110-Al.

XX 23-OCT-1997.

PF 15-APR-1997; 97WO-US06300.

PR 16-APR-1996; 96US-0015517.

XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.

PA (WISC) WISCONSIN ALMONI RES FOUND.

XX Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;

XX WPI: 1997-526447/48.

XX N-PSDB: T93293.

XX Transgenic plants expressing geminivirus AC1 and Cl wild-type and

mutant genes have increased resistance to geminivirus infection

PT e-9 tomato mottle virus and tomato yellow leaf curl virus or bean

PT golden mosaic geminivirus

XX Example 5; Page 115-116; 132pp; English.

XX This protein comprises a transdominant lethal mutant of the Bean

golden mosaic virus (BGWV) Cl protein (see W34338) that is required

for replication. It is encoded by mutated Cl open reading frame

CC BGAC262 (see T93293) and carries a mutation in the NTP-binding

CC domain. The invention involves production of transgenic plants

CC containing DNA comprising geminivirus Cl or AC1 wild-type or mutant

CC sequences that negatively interfere in trans with geminiviral

CC replication during infection. Such transgenic plants are resistant

CC to viral infection. The AC1/Cl genes are especially from BGWV.

CC tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)

CC and encode polypeptides (see W34324-35) that have mutations in the

CC highly conserved DNA-nicking and/or the NTP-binding domains.

XX Sequence 353 AA;

Query Match 61.3%; Score 222; DB 18; Length 353;

Best Local Similarity 60.0%; Pred. No. 7.2e-20;

Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

Oy 1 TLWGEFVQAGRSAGCCTSDAAAEALNASSKEALQIIRKIPKYLQFHNLSNL 60

Db 110 TLWGEFVQAGRSAGCCTSDAAAEALNASSKEALQIIRKIPKYLQFHNLSNL 169

QY 61 DRFDKTPP 70  
 Db 170 elivavppap 179

RESULT 6  
 ID R88870 standard; Protein: 359 AA.

XX R88870;  
 XX  
 XX R88870;  
 XX  
 XX 07-NOV-1996 (first entry)  
 XX  
 XX Sardinian tomato yellow leaf curl virus mutated Rep protein (K227A).  
 XX Nucleotide triphosphate binding site; DNA helicase; RNA helicase;  
 KW modification; mutation; viral replication; deficient; inhibition;  
 KW viral resistance; geminivirus; tomato yellow leaf curl virus;  
 KW Sardinian isolate; STYLICV; transgenic plant; P-loop; C1 protein;  
 KW A11 protein; dominant negative phenotype.  
 XX Sardinian tomato yellow leaf curl virus.

XX Key Location/Qualifiers  
 FT Misc-difference 227  
 PT /note= "wild-type Lys has been replaced by Ala"  
 XX  
 PN W09608573-A1..

XX 21-MAR-1996.  
 XX 15-SEP-1995; 95WO-FR01192.  
 XX 15-SEP-1994; 94FR-0011040.  
 XX (CNRS ) CENT NAT RECH SCI.  
 XX Gronenborn B;

XX WPI: 1996-179947/18.  
 DR N-PSDB: T12904.

PT Prodn. of virus-resistant transgenic plants - using mutated genomic  
 PT sequence from phytopathogenic DNA virus  
 XX  
 PS Disclosure: Fig 13; 93pp; French.

XX Mutation of consensus amino acids in the NTP-binding site of  
 CC geminivirus Rep protein is used to produce replication deficient  
 CC viruses. The mutated viral nucleic acid is used for producing  
 CC transgenic plants that are resistant to, or tolerant of, the native  
 CC virus. The present sequence is a mutant form of the Rep (or C1)  
 CC protein from the Sardinian isolate of tomato yellow leaf curl virus  
 CC (STYLICV) in which the wild-type Lys227 residue has been changed to an  
 CC Ala residue; transgenic Nicotiana benthamiana plants generated by  
 CC transformation with the mutated plants were found to be resistant to  
 CC STYLICV, i.e. the mutation results in a dominant negative phenotype.

XX Sequence 359 AA;

Query Match 61.3%; Score 222; DB 17; Length 359;  
 Best Local Similarity 58.0%; Pred. No. 7.3e-20;  
 Matches 40; Conservative 11; Mismatches 18; Indels 0; Gaps 0;

QY 2 LVWGEFQVAGRSAGCGCOTSDNDAAEALNASKEEALQITIREKTYLQFPHNLNSLD 61  
 Db 111 lewgtftqdgrragggqandayakainagsqaldvikelapdyvlhfhlnsld 170

QY 62 RFDKTPP 70

Db 171 kvfqvppap 179

RESULT 7

XX R88870 standard; Protein: 359 AA.

XX R88870;  
 XX  
 XX 07-NOV-1996 (first entry)

XX Sardinian tomato yellow leaf curl virus mutated Rep protein (K227H).  
 XX Nucleotide triphosphate binding site; DNA helicase; RNA helicase;  
 KW modification; mutation; viral replication; deficient; inhibition;  
 KW viral resistance; geminivirus; tomato yellow leaf curl virus;  
 KW Sardinian isolate; STYLICV; transgenic plant; P-loop; C1 protein;  
 KW A11 protein.

XX Sardinian tomato yellow leaf curl virus.

XX Key Location/Qualifiers  
 FT Misc-difference 227  
 PT /note= "wild-type Lys has been replaced by His"

XX W09608573-A1.

XX 21-MAR-1996.  
 XX 15-SEP-1995; 95WO-FR01192.  
 XX 15-SEP-1994; 94FR-0011040.

XX (CNRS ) CENT NAT RECH SCI.

XX Gronenborn B;

XX WPI: 1996-179947/18.  
 DR N-PSDB: T12905.

PT Prodn. of virus-resistant transgenic plants - using mutated genomic  
 PT sequence from phytopathogenic DNA virus

XX Disclosure: Fig 13; 93pp; French.

XX Mutation of consensus amino acids in the NTP-binding site of  
 CC geminivirus Rep protein is used to produce replication deficient  
 CC viruses. The mutated viral nucleic acid is used for producing  
 CC transgenic plants that are resistant to, or tolerant of, the native  
 CC virus. The present sequence is a mutant form of the Rep (or C1)  
 CC protein from the Sardinian isolate of tomato yellow leaf curl virus  
 CC (STYLICV) in which the wild-type Lys227 residue has been changed to a  
 CC His residue; transgenic Nicotiana benthamiana plants generated by  
 CC transformation with the mutated plants were not resistant to STYLICV.  
 CC In contrast, plants transformed with a virus in which Lys227 had been  
 CC replaced by Ala were found to be resistant.

XX Sequence 359 AA;

Query Match 61.3%; Score 222; DB 17; Length 359;  
 Best Local Similarity 58.0%; Pred. No. 7.3e-20;  
 Matches 40; Conservative 11; Mismatches 18; Indels 0; Gaps 0;

QY 2 LVWGEFQVAGRSAGCGCOTSDNDAAEALNASKEEALQITIREKTYLQFPHNLNSLD 61  
 Db 111 lewgtftqdgrragggqandayakainagsqaldvikelapdyvlhfhlnsld 170

QY 62 RFDKTPP 70

Db 171 kvfqvppap 179

RESULT 8

```

R88872
ID R88872 standard; Protein; 359 AA.
AC
XX R88872;
XX
DT 07-NOV-1996 (first entry)
XX
DE Sardinian tomato yellow leaf curl virus mutated Rep protein (K227R).
XX
XX Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
XX modification; mutation; viral replication; deficient; inhibition;
XX RNA resistance; geminivirus; tomato yellow leaf curl virus;
XX Sardinian isolate; STYLCV; transgenic plant; P-loop; CI protein;
XX ALI protein.
XX
XX Sardinian tomato yellow leaf curl virus.
XX
XX
XX Key Location/Qualifiers
XX Misc-difference 227
XX /note= "wild-type Lys has been replaced by Arg"
XX
XX W09608573-A1.
XX
XX 21-MAR-1996.
XX
XX 15-SEP-1995; 95WO-FR01192.
XX
XX 15-SEP-1994; 94FR-0011040.
XX
XX (CNRS ) CENT NAT RECH SCI.
XX
XX Gronsenborn B;
XX
XX WPI: 1996-179947/18.
XX
XX N-PSDB; 112906.
XX
XX Prodn. of virus-resistant transgenic plants - using mutated genomic
XX sequence from phytopathogenic DNA virus
XX
XX Disclosure; Fig 13; 93pp; French.
XX
XX Mutation of consensus amino acids in the MTP-binding site of
XX geminivirus Rep protein is used to produce replication deficient
XX transgenic plants that are resistant to or tolerant of the same
XX virus. The present sequence is a mutant form of the Rep (or CI)
XX protein from the Sardinian isolate of tomato yellow leaf curl virus
XX (STYLCV) in which the wild-type Lys227 residue has been changed to
XX an Arg residue; transgenic Nicotiana benthamiana plants generated by
XX transformation with the mutated virus were not resistant to STYLCV.
XX CC transgenic flowers transformed with a virus in which Lys227 had been
XX replaced by Ala were found to be resistant.
XX
XX Sequence 359 AA:

Query Match 61.3%; Score 222; DB 17; Length 359;
Best Local Similarity 59.04; Fred; No. 7.3e-20;
Matches 40; Conservative 11; Mismatches 18; Indels 0; Gaps 0;
Oy 2 LWNKEFQVAGRSARGCGTSDNDAAEALNASSKEEALQIREKIPKYLFOFHINSNLD 61
Db 111 lllllllllllllllllllllllllllllllllllllllllllllllllllllll
Oy 62 RIFQKTPPEP 70
Db 171 kvrqqvpep 179

RESULT 9
P70407
ID P70407 standard; Protein; 353 AA.
XX

```

```

P70407:
AC
XX 02-MAY-1991 (first entry)
XX
XX ORF 4 gene product of Bean Golden Mosaic virus.
XX
XX Plant vector.
XX
XX Bean golden mosaic virus.
XX
XX JP61257186-A.
XX
XX 14-NOV-1986.
XX
XX 10-MAY-1985; 85JP-0098108.
XX
XX 10-MAY-1985; 85JP-0098108.
XX
XX (TEIJ ) TEIJOIN KK.
XX
XX WPI: 1987-159662/23.
XX
XX N-PSDB; N70630.
XX
XX New DNA and hybrid DNA - used for recombinant vector of plants.
XX
XX Disclosure; Fig 6; 24pp; Japanese.
XX
XX The sequence encoding this protein may be taken along with the -a
XX DNA fragment and used to create a stable resistance gene, and used to create a
XX recombinant plant vector.
XX
XX See also N70629.
XX
XX Sequence 353 AA:

Query Match 50.8%; Score 220; DB 8; Length 353;
Best Local Similarity 50.04; Fred; No. 1.2e-16;
Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;
Oy 1 TLWNGEFQVAGRSARGCGTSDNDAAEALNASSKEEALQIREKIPKYLFOFHINSNLD 60
Db 110 tiewgfgvqgrsargqgsandsykalnadsesaitllkeepkdyvqhnlrsl 169
Oy 61 DRIFQKTPPEP 70
Db 170 eriffvpep 179

RESULT 10
W34336
ID W34336 standard; Protein; 361 AA.
XX
XX W34336;
XX
XX 27-APR-1998 (first entry)
XX
XX Tomato mottle virus AC1 protein.
XX
XX Geminivirus; TOMOV; AC1 gene; transdominant mutation;
XX transgenic plant; disease resistance.
XX
XX Tomato mottle virus isolate Florida.
XX
XX W09739110-A1.
XX
XX 23-OCT-1997.
XX
XX 15-APR-1997; 97MO-US05300.
XX
XX 16-APR-1996; 96US-0015517.
XX
XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.
XX
XX (WISC ) WISCONSIN ALUMNI RES FOUND.
XX

```

XX	PI	Ahliquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX	DR	WPI; 1997-526447/48..
XX	DR	N-PSDB; T93294.
XX	PT	Transgenic plants expressing geminivirus AC1 and CI wild-type and
XX	PT	mutant genes that increased resistance to geminivirus infection
XX	PT	in tomato. Genes for the wild-type and mutant AC1 and CI genes
XX	PT	from the yellow leaf curl virus of bean
XX	PT	golden mosaic geminivirus
XX	PT	
XX	PS	Example 3.3; Page 57-58; 132pp; English.
XX	PS	
XX	CC	This protein comprises the wild-type AC1 protein of tomato mottle
XX	CC	virus (TOMOV), a geminivirus that has a bipartite genome. The AC1
XX	CC	gene (see T93294) must be expressed for efficient replication of
XX	CC	the two genomic components DNA-1 and DNA-2. The AC1 protein has
XX	CC	DNA binding activity, DNA nicking activity, and DNA nicking
XX	CC	activity, and an NTP binding activity. The invention involves
XX	CC	production of transgenic plants containing DNA comprising AC1 or CI
XX	CC	wild-type or mutant sequences that negatively interfere in trans
XX	CC	with geminiviral replication during infection. Such transgenic
XX	CC	plants are resistant to viral infection. The AC1/CI genes are
XX	CC	especially from TOMOV, tomato yellow leaf curl virus or bean golden
XX	CC	mosaic geminivirus (see T93282-93) and encode polypeptides (see
XX	CC	T93284-35) that have mutations in the highly conserved DNA-nicking
XX	CC	domain and/or the NTP-binding domain.
XX	CC	
XX	CC	Sequence 361 AA;
XX	SQ	

	Query Match	Best Local Similarity	60.8%	Score 220	DB 18	Length 361
	Matches 40	Conservative 16	Mismatches 14	Indels	0	Gaps
Qy	1	TLWNGFEVQVARGSGQTSNDAAALANASSKEALQIRKIPKVLQFQFINLSNL	60			
Db	110	tlwngfdlqrsrsgqgaqdaakalanassvsgalavireepkdvqlqnhairsl	169			
Qy	61	DRITWDTEP 70				
Db	170	erllkanep 179				

RESULT 11	
K34324	
ID	W34324 standard; Protein; 361 AA.
DE	XX
DT	W34324;
DD	XX
DT	27-APR-1998 (first entry)
DE	XX
DE	Tomato mottle virus AC1 mutant TOMV-AC1dLn.
DE	XX
DE	Geminivirus: TOMOV-AC1dLn; AC1 gene; transdominant mutation;
DE	transgenic plant; disease resistance.
XX	XX
XX	Tomato mottle virus isolate Florida.
OS	Synthetic.
XX	XX
PN	W09739110-AL.
XX	XX
XX	23-OCT-1997.
XX	XX
PF	15-APR-1997; 97MO-US06300.
XX	XX
PR	16-APR-1996; 96US-0015517.
XX	XX
XX	(SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA	(WISC) WISCONSIN ALUMINI RES FOUND.
PA	XX
XX	Abiquist PG, Hanson SE, Luu HT, Maxwell DP, Stout JT:
XX	

DR	WPI: 1997-526447/48.
PT	N-PSDB: T93282.
PT	
PT	Transgenic plants expressing geminivirus AC1 and CI wild-type and
PT	mutant genes - have increased resistance to geminivirus infection
PT	e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT	golden mosaic geminivirus
XX	
XX	
PS	Example 3.4; Page 60-62; 132pp; English.
CC	
CC	This protein comprises a transdominant lethal mutant, designated
CC	Tomov-AC1dm, of tomato mottle virus (Tomov) AC1 protein (see
CC	W43436). It is encoded by a mutant AC1 gene (see T93282) of
CC	tomato mottle virus (Tomov), and carries 3 mutations in its
CC	NTP-binding domains. The AC1 gene (see also T93294) must be
CC	expressed in a highly efficient replication of the two genomic components,
CC	involves packaging of transgenic plants containing DNA comprising
CC	geminivirus AC1 or CI wild-type or mutant sequences that negatively
CC	interfere in trans with geminiviral replication during infection. The
CC	such transgenic plants are resistant to viral infection. The
CC	AC1/CI genes are especially from Tomov, tomato yellow leaf curl
CC	virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC	polypeptides (see W43424-35) that have mutations in the highly
CC	conserved DNA-ACKING and/or NTP-binding domains.
XX	
XX	Sequence 361 AA;
XX	

	Query Match	60.9%	Score 220	DB 18	Length 361;
	Best Local Similarity	57.1%	Pred No. 1.3e-19;		
	Matches 40:	Conservative 16;	Mismatches 14;	Indels 0;	Gaps
Qy	1 TLVGFGEFVGARSARGCGTSDNDAALNNASKEALQIRKIPKYLVQFHNLNSL 60				
Dy	110 TLEGDFGLDGRSGRGGAANDSYAKA HANSSGVSLAVIECPKDVQLGHNRSL 169				
	:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :				
Qy	51 DRIFKWTFEP 70				
Dy	51 DRIFKWTFEP 70				
Qy	170 ERIFKANRP 179				
Dy	170 ERIFKANRP 179				

RESULT 12	
W34325	W34325 standard; Protein; 361 AA.
XX	
AC	W34325;
XX	
DT	27-APR-1998 (first entry)
XX	
DE	Tomato mottle virus ACl mutant TOWV-ACldm1.
XX	
DE	Geminivirus: TOWV-ACldm1; ACl gene; transdominant mutation;
XX	
AC	transgenic plant; disease resistance.
XX	
KW	Tomato mottle virus isolate Florida.
XX	
OS	Synthetic.
XX	
PN	W09739110-A1.
XX	
XX	23-OCT-1997.
XX	
PF	15-APR-1997; 97WO-US06300.
XX	
XX	16-APR-1996; 96US-0015517.
XX	
PA	(SEMI- } SEMINIS VEGETABLE SEEDS INC.
PA	(WISC } WISCONSIN ALUMNI RES FOUND.
XX	
PI	Ahquist PG, Hanson SF, Lou HT, Maxwell DP, Scout JT.
XX	
DR	WPI: 1997-526447/48.
DR	N-PSDB: T93283.

XX Transgenic plants expressing geminivirus AC1 and CI wild-type and  
 PT mutant genes - have increased resistance to geminivirus infection  
 PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean  
 PT golden mosaic geminivirus  
 XX  
 PS Example 3.5; Page 64-65; 132pp; English.  
 XX  
 XX This protein comprises a transdominant lethal mutant, designated  
 CC TOMoV-AC1dm23, of tomato mottle virus (TOMoV) AC1 protein (see  
 CC W34326). It is encoded by a mutant AC1 gene (see T93284) of  
 CC tomato mottle virus (TOMoV), and carries a mutation in an  
 CC NTP-binding domain. The AC1 gene (see also T93294) must be  
 CC expressed for efficient replication of the two genomic components,  
 CC DNA-A and DNA-B, of the bipartite TOMoV genome. The DNA  
 CC involves production of transgenic plants containing DNA  
 CC interfering in trans with geminiviral replication during infection.  
 CC Such transgenic plants are resistant to viral infection. The  
 CC AC1/CI genes are especially from TOMoV, tomato yellow leaf curl  
 CC virus or bean golden mosaic geminivirus (see T93282-93) and encode  
 CC polypeptides (see W3424-35) that have mutations in the highly  
 CC conserved DNA-nicking and/or NTP-binding domains.  
 XX Sequence 361 AA;  
 SQ  
 Query Match 60.8%; Score 220; DB 18; Length 361;  
 Best Local Similarity 57.1%; Pred. No. 1.3e-19;  
 Matches 40; Conservative 16; Mismatches 14; Indels 0; Gaps 0;  
 QY 1 TLVGEFVQAGRSAGCGCTSNDAAEALNASSKEEAQIIREKTPKYLQFHNLNSLI 60  
 DB 110 TLVGEFVQAGRSAGCGCTSNDAAEALNASSKEEAQIIREKTPKYLQFHNLNSLI 60  
 QY 61 DRIEDKTPPEP 70  
 DB 170 erifakaep 179  
 RESULT 14  
 W34326 ID P70562 standard; Protein: 361 AA.  
 AC AC P70562;  
 XX AC W34326;  
 DT 27-APR-1998 (first entry)  
 XX Tomato mottle virus AC1 mutant TOMV-AC1dm23.  
 XX Geminivirus; TOMoV-AC1dm23; AC1 gene; transdominant mutation;  
 KW Transgenic plant; disease resistance.  
 XX  
 OS Tomato mottle virus isolate Florida.  
 QS Synthetic.  
 XX  
 PN W05739110-AI.  
 XX 23-OCT-1997.  
 XD  
 XX 15-APR-1997; 97NO-US06300.  
 PF  
 XX 16-APR-1996; 96US-0015517.  
 PR  
 XX (SMT-) SEMINIS VEGETABLE SEEDS INC.  
 PA (WISC ) WISCONSIN ALUMNI RES FOUND.  
 XX  
 XX Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;  
 WPI: 1997-526447/48.  
 DR N-PSDB; T93284.6NP.  
 XX Transgenic plants expressing geminivirus AC1 and CI wild-type and

PT mutant genes - have increased resistance to geminivirus infection  
 PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean  
 PT golden mosaic geminivirus  
 XX  
 PS Example 3.6; Page 67-69; 132pp; English.  
 XX  
 XX This protein comprises a transdominant lethal mutant, designated  
 CC TOMoV-AC1dm23, of tomato mottle virus (TOMoV) AC1 protein (see  
 CC W34326). It is encoded by a mutant AC1 gene (see T93284) of  
 CC tomato mottle virus (TOMoV), and carries 2 mutations in an  
 CC NTP-binding domain. The AC1 gene (see also T93294) must be  
 CC expressed for efficient replication of the two genomic components,  
 CC DNA-A and DNA-B, of the bipartite tomato mottle virus genome.  
 CC involves production of transgenic plants containing DNA  
 CC interfering in trans with geminiviral replication during infection.  
 CC Such transgenic plants are resistant to viral infection. The  
 CC AC1/CI genes are especially from TOMoV, tomato yellow leaf curl  
 CC virus or bean golden mosaic geminivirus (see T93282-93) and encode  
 CC polypeptides (see W3424-35) that have mutations in the highly  
 CC conserved DNA-nicking and/or NTP-binding domains.  
 XX Sequence 361 AA;  
 SQ  
 Query Match 60.8%; Score 220; DB 18; Length 361;  
 Best Local Similarity 57.1%; Pred. No. 1.3e-19;  
 Matches 40; Conservative 16; Mismatches 14; Indels 0; Gaps 0;  
 QY 1 TLVGEFVQAGRSAGCGCTSNDAAEALNASSKEEAQIIREKTPKYLQFHNLNSLI 60  
 DB 110 TLVGEFVQAGRSAGCGCTSNDAAEALNASSKEEAQIIREKTPKYLQFHNLNSLI 60  
 QY 61 DRIEDKTPPEP 70  
 DB 170 erifakaep 179  
 RESULT 14  
 P70562 ID P70562 standard; Protein: 361 AA.  
 AC AC P70562;  
 XX AC 30-APR-1991 (first entry)  
 XX Product of ORF 4 from MYMV complementary strand (b).  
 DE  
 XX Geminivirus.  
 KW  
 XX Mungbean yellow mosaic virus.  
 OS  
 XX P62126982-A.  
 XX  
 PD 09-JUN-1987.  
 XX  
 PF 28-NOV-1985; 85JP-0266080.  
 XX  
 PR 28-NOV-1985; 85JP-0266080.  
 XX  
 PA (TEIJ ) TEIJIN KK.  
 XX  
 DR WPI: 1987-196308/28.  
 XX N-PSDB; N70895.  
 XX Novel DNA and hybrid DNA useful - as vector for recombinant work  
 PT of plant gene.  
 PT Disclosure; Fig 8; 21pp; Japanese.  
 XX The sequence is encoded by ORF 4 which occurs on the complementary  
 CC strand of the (b) molecule of the geminivirus.  
 CC See also P70559-P70567.

Sat Feb 3 10:18:52 2001

us-09-289-346a-9.rag

```

XX SQ Sequence 361 AA;
Query Match 59.0%; Score 213.5; DB 8; Length 361;
Best Local Similarity 66.2%; Pred No. 8; e-19;
Matches 43; Conservative 9; Mismatches 12; Indels 1; Gaps 1;

QY 5 GEFQVAGRSARGCOTSDNDAAEALNASSKEEALQIIREKIPKYLQFQHNLSNLDRI 64
DQ 114 gsfqvdgrsrgskqandayaeaingsgkqlalnllrekapkeyllqfhnlnslrif 173
QY 65 -DKTP 68
DQ 174 sdevp 178

RESULT 15
W56495 W56495 standard; Protein; 362 AA.
XX AC W56495;
XX DT 11-AUG-1998 (first entry)
XX DE Tobacco leaf curling virus gene protein Cl.
XX KX Tobacco leaf curling virus gene; TlCV; promoter; Cl protein.
XX OS Tobacco leaf curling virus.
XX PN JF10070982-A.
XX QD 17-MAR-1998.
XX PR 30-AUG-1996; 96JP-0230394.
XX PR 30-AUG-1996; 96JP-0230394.
XX PA (NORQ) NORINSUISANSHO KYUSHU NOGYO SHIKENJYO.
XX WP1; 1998-233630/21.
XX PT Tobacco leaf curling virus gene - useful for inserting into vectors
XX PT for expression in, e.g. tomato plants.
XX PS Claim 5; Fig 5; 9pp; Japanese.
XX CC This sequence represents the Cl protein encoded by the tobacco leaf
XX CC curling virus (TlCV) gene of the invention. TlCV gene or its promoter can
XX CC be inserted into a vector for expression in plants, e.g. tobacco and
XX CC tomato. This sequence is believed to be encoded by the TlCV gene shown in
XX CC V29761.
XX SQ Sequence 362 AA;
Query Match 57.5%; Score 208; DB 19; Length 362;
Best Local Similarity 57.4%; Pred No. 4; e-14;
Matches 44; Conservative 10; Mismatches 14; Indels 16; Gaps 1;

QY 2 LWGGEFQVAGRSARGCOTSDNDAAEALNASSKEEALQIIREKIPKYLQFQHNLSNLD 61
DQ 111 ldfgfvqdrsrccgcsandayaeaingsgksaalnllrekapkdrvlqfhnlnsl 170
QY 62 RT-----DKTP 69
DQ 171 rlfaplelvfcvfssssfdqpye 194

```

Search completed: February 3, 2001, 02:15:30  
 Job time: 3437 sec



GenCore version 4.5  
 Copyright (c) 1993 - 2000 CompuGen Ltd.  
 OM protein - protein search, using sw model  
 Run on: February 3, 2001, 02:22:51 : Search time 180.33 Seconds  
 (without alignments)  
 45.497 Million cell updates/sec

Title: US-09-289-346a-7  
 Perfect score: 359  
 Sequence: 1 TLVWGEFVDGRSGRGCGT.....FOFHNLSALAIIDFTPEP 70

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPRTEMBL.15.\*

1: sp.archaea.\*

2: sp.bacteria.\*

3: sp.fungi.\*

4: sp.human.\*

5: sp.invertebrate.\*

6: sp.plant.\*

7: sp.vertebrate.\*

8: sp.organelle.\*

9: sp.phage.\*

10: sp.plant.\*

11: sp.rodent.\*

12: sp.virus.\*

13: sp.vertebrate.\*

14: sp.unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	274	76.3	226	12	O09727 leonurus mo
2	274	76.3	226	12	O9WHF6 tomato mild
3	272	75.8	361	12	O67574 bean golden
4	270	75.2	225	12	O9QDB1 cowpea gold
5	269	74.8	225	12	O9QDB1 cowpea gold
6	269	74.8	185	12	O9QDB1 cowpea gold
7	259	72.1	149	12	P88975 macropittu
8	259	72.1	233	12	O9YIA4 macropittu
9	258	71.9	234	12	O39180 potato yell
10	252	70.2	190	12	O9Z089 tobacco lea
11	252	70.2	190	12	O9Z084 tobacco lea
12	249	69.4	190	12	O9WH27 tobacco lea
13	249	69.4	253	12	O9Z084 tobacco lea
14	249	69.4	253	12	O9Z084 tobacco lea
15	249	69.4	354	12	O65438 beet curly
16	247	68.8	208	12	O9Z0C0 tobacco lea
17	247	68.8	208	12	O9Z0B8 tobacco lea
18	245	68.5	203	12	O9Z083 tobacco lea
19	245	68.2	208	12	O9Z0B6 tobacco lea

20	245	68.2	358	12	O65418	065418 beet curly
21	244	68.0	363	12	O73577	073577 cotton leaf
22	238	66.3	190	12	O9Z086	09Z086 tobacco lea
23	238	66.3	362	12	O58816	058816 chayote mos
24	237	66.0	362	12	O9Z086	09Z086 tobacco lea
25	237	66.0	362	12	O9Z086	09Z086 tobacco lea
26	237	66.0	363	12	O72705	072705 cotton leaf
27	237	66.0	363	12	O72719	072719 cotton leaf
28	236	65.7	208	12	O9Z0C6	09Z0C6 tobacco lea
29	235	65.5	349	12	O88888	088888 tomato pseu
30	234	65.2	359	12	O9Z2V4	09Z2V4 tomato yell
31	234	65.2	359	12	O9Z2V2	09Z2V2 tomato yell
32	234	65.2	359	12	O9Z2V2	09Z2V2 tomato yell
33	234	65.2	359	12	O9Z2V2	09Z2V2 tomato yell
34	233	64.9	359	12	O88942	088942 tobacco lea
35	232	64.6	190	12	O9Z0A7	09Z0A7 tobacco lea
36	231	64.3	307	12	O91ET7	091ET7 cotton leaf
37	231	64.3	361	12	O72723	072723 cotton leaf
38	230	64.1	231	12	O95620	095620 african tom
39	230	64.1	231	12	O91N42	091N42 south afric
40	229	63.8	358	12	O95620	095620 african tom
41	229	63.8	358	12	O91E42	091E42 cassava gen
42	228	63.5	358	12	O9WR17	09WR17 african cas
43	228	63.5	360	12	O9YNW5	09YNW5 sida golden
44	226	63.0	351	12	O9Q9R3	09Q9R3 dicliptera
45	225	62.7	362	12	O9YIT7	09YIT7 althea rose

## ALIGNMENTS

RESULT 1  
 ID O09727 PRELIMINARY: PRT; 226 AA.  
 AC O09727: 1957 (TEMBREL. 04, Created)  
 DT 01-JUL-1997 (TEMBREL. 04, Last sequence update)  
 DT 01-OCT-2000 (TEMBREL. 15, Last annotation update)  
 DE REP-PROTEIN (FRAGMENT).  
 GN REP.  
 OS Leonurus mosaic virus.  
 OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=58177;  
 RN 1  
 RC SEQUENCE FROM N.A.  
 RA Faria J.C., Maxwell D.P.;  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DOJB databases.  
 DR EMBL: U92532; AAB51157.1;  
 DR INTERPRO: IPR001191;  
 DR INTERPRO: IPR001301;  
 DR PRINTS: PR00222; GEMCATAL1.  
 DR PRINTS: PR00222; GEMCATAL1.  
 FT NON\_TER 226 226  
 SQ SEQUENCE 226 AA; 25617 MW; 73CDB6E766083FC5 CRC64;

Query Match  
 Percent Similarity 76.3%; Score 274; DB 12; Length 226;  
 Matches 54; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

Oy 1 TLVWGEFVDGRSGRGCGTGNDAALANSSKEALQIIRKIPKYLFOFHNLSAL 60  
 Db 111 TVWGEFVDGRSGRGCGTGNDAALANSSKEALQIIRKIPKYLFOFHNLSAL 170

Oy 61 AAIIDFTPEP 70

RESULT 2  
 O9WHF6

```

ID Q9WHF6 PRELIMINARY; PRT: 225 AA.
AC Q9WHF6
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN REP.
OS tomato mild mottle geminivirus.
OC viruses: ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.
NCBI_TaxID=32943;
RN [1] JENSEN FROM N.A.
RC STRAIN=BN56-HSKA.
RA Nakhla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,
RA Maxwell D.P.;
RT "Molecular characterization and DNA-based detection methods for
RT vegetable-infecting geminiviruses in Central America.";
RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF131071; AAD33471.1;
DR INTERPRO: IPR001301;
DR PFAM: PF00759; GeminLALL; 1.
DR PRINTS: PR00227; GEMCOATALL.
DR PRINTS: PR00228; GEMCOATCLVL.
DR NON-TER 226 226
FT SEQUENCE 226 AA; 25941 MW; 2E4116712871A23 CRC64;
SQ

Query Match 76.3%; Score 274; DB 12; Length 226;
Best Local Similarity 72.3%; Pred. No. 2e-23;
Matches 51; Conservative 12; Mismatches 7; Indels 0; Gaps 0;

QY 1 TLWGFQVDSRGSGCOTSDNAAALNASSKEALQIREKIPEKYLFOFHNLSAL 60
DB 111 TIEWGFQDGRSARGGQOTANDAAASEALNASSKEAQIIEKLPKFLFOFHNLSNL 170
QY 61 AATFKTPEP 70
DB 171 DRIFKAPEP 180

RESULT 3
ID Q6757A PRELIMINARY; PRT: 361 AA.
AC Q6757A
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)
DE POTATIVE REPLICATIVE PROTEIN.
GN AC1.
OS A.L. golden mosaic virus.
OC viruses: ssDNA viruses; Geminiviridae; Begomovirus.
NCBI_TaxID=10839;
RN [1]
RT SEQUENCE FROM N.A.
RA Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,
RA Maxwell D.P., Russell D.R.;
RL "Molecular characterization and DNA-based detection methods for
RL vegetable-infecting geminiviruses in Central America.";
RN Submitted (MAY-1992) to the EMBL/Genbank/DBJ databases.
DR EMBL: M86586; AAA46312.1;
DR INTERPRO: IPR001191;
DR PFAM: PF00759; GeminLALL; 1.
DR PRINTS: PR00227; GEMCOATALL.
DR PRINTS: PR00228; GEMCOATCLVL.
SQ

Query Match 75.2%; Score 270; DB 12; Length 225;
Best Local Similarity 76.1%; Pred. No. 5.7e-23;
Matches 51; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 4 WGFQVDSRGSGCOTSDNAAALNASSKEALQIREKIPEKYLFOFHNLSALAAI 63
DB 113 WGFQVDSRGSGCOTSDNAAALNASSKEAQIIEKLPKFLFOFHNLSNLDRI 172
QY 64 FDKTPEP 70
DB 173 FRAPPEP 179

RESULT 5
ID Q9QS55 PRELIMINARY; PRT: 364 AA.
AC Q9QS55
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DE REPLICATION INITIATION PROTEIN AC1.
GN AC1.
OS sweet potato leaf curl virus.
OC viruses: ssDNA viruses; Geminiviridae; Begomovirus.
NCBI_TaxID=100755;
RN [1]
RT SEQUENCE FROM N.A.

```

```

ID PD000736; -. 1.
AC PD000736; -. 1.
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN REP.
OS corn golden mosaic geminivirus.
OC viruses: ssDNA viruses; Geminiviridae; Begomovirus.
NCBI_TaxID=63469;
RN [1]
RT SEQUENCE FROM N.A.
RC STRAIN=CGNV-BR.
RA Faria J.C.;
RT "Partial nucleotide sequence of corn golden mosaic geminivirus from
RT Brazil, (SEP-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF168708; AAF06318.1;
DR INTERPRO: IPR001301;
DR PFAM: PF00759; GeminLALL; 1.
DR PRINTS: PR00227; GEMCOATALL.
DR PRINTS: PR00228; GEMCOATCLVL.
DR NON-TER 225 225
FT SEQUENCE 225 AA; 23766 MW; 1089CB6BD8D15B5D CRC64;
SQ

Query Match 75.2%; Score 270; DB 12; Length 225;
Best Local Similarity 76.1%; Pred. No. 5.7e-23;
Matches 51; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 4 WGFQVDSRGSGCOTSDNAAALNASSKEALQIREKIPEKYLFOFHNLSALAAI 63
DB 113 WGFQVDSRGSGCOTSDNAAALNASSKEAQIIEKLPKFLFOFHNLSNLDRI 172
QY 64 FDKTPEP 70
DB 173 FRAPPEP 179

RESULT 5
ID Q9QS55 PRELIMINARY; PRT: 364 AA.
AC Q9QS55
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DE REPLICATION INITIATION PROTEIN AC1.
GN AC1.
OS sweet potato leaf curl virus.
OC viruses: ssDNA viruses; Geminiviridae; Begomovirus.
NCBI_TaxID=100755;
RN [1]
RT SEQUENCE FROM N.A.

```

RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;  
 RT "Detection of a geminivirus infecting sweet potato in the United  
 States"; 82:1253-1257(1998).  
 RL Plant Dis. 82:1253-1257(1998).  
 RC INTERPRO: IP0001191; -  
 DR INTERPRO: IP0001191; -  
 DR PFAM: PF00759; Gemini\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATL1.  
 SQ SEQUENCE 364 AA; 40680 MW; 5F75752431A09D6E CRC64;

Query Match 75.28; Score 270; DB 12; Length 364;  
 Best Local Similarity 79.48; Pred. No. 9.8e-23;  
 Matches 54; Conservative 3; Mismatches 11; Indels 0; Gaps 0;  
 Oy 1 TLWGFEVDGRSARGCOTSDAAAEALNASKEEALQIIRKIPKYLFOFHNLNSAI 60  
 Db 110 TLWGFEVDGRSARGCOTSDAAAEALNASKEEALQIIRKIPKYLFOFHNLNSAI 169  
 Oy 61 AAFDKTP 68  
 Db 170 DRIESPP 177

RESULT 6  
 ID Q96933 PRELIMINARY; PRT; 185 AA.  
 AC Q96933;  
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE REP PROTEIN (FRAGMENT).  
 GN ACL.  
 OS "A golden mosaic virus".  
 OC VIRUS: SMDNA viruses, Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=51034;  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-JAMAICA;  
 RA Roye M.E., McLaughlin W.A., Nekhla N.K., Maxwell D.P.;  
 RL Plant Dis. 81:1251-1258(1997).  
 RC INTERPRO: IP0001191; -  
 DR INTERPRO: IP0001191; -  
 DR PFAM: PF00759; Gemini\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATL1.  
 FT NON\_TER 185  
 FT NON\_TER 185  
 SQ SEQUENCE 185 AA; 20975 MW; 3913850A02545EE1 CRC64;

Query Match 74.48; Score 267; DB 12; Length 185;  
 Best Local Similarity 70.08; Pred. No. 9.9e-23;  
 Matches 49; Conservative 12; Mismatches 9; Indels 0; Gaps 0;  
 Oy 1 TLWGFEVDGRSARGCOTSDAAAEALNASKEEALQIIRKIPKYLFOFHNLNSAI 60  
 Db 89 TLWGFEVDGRSARGCOTSDAAAEALNASKEEALQIIRKIPKYLFOFHNLNSAI 148  
 Oy 61 AAFDKTP 70  
 Db 149 DRFSAPP 158

RESULT 7  
 ID P88975 PRELIMINARY; PRT; 149 AA.  
 AC P88975;  
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)

DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).  
 GN ACL.  
 OS Macropitium golden mosaic geminivirus.  
 OS VIRUSES: SMDNA viruses, Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=51676;  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-JAMAICA;  
 RA Roye M.E.;  
 RL Thesis (1996). Biochemistry, University of the West Indies, Jamaica.  
 DR EMBL: U75278; A0356919.1; -  
 DR INTERPRO: IP0001191; -  
 DR INTERPRO: IP0001191; -  
 DR PFAM: PF00759; Gemini\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATL1.  
 FT NON\_TER 149  
 FT NON\_TER 149  
 SQ SEQUENCE 149 AA; 16785 MW; E4CF5EED4C9CD508 CRC64;

Query Match 72.18; Score 259; DB 12; Length 149;  
 Best Local Similarity 68.68; Pred. No. 6.3e-22;  
 Matches 48; Conservative 12; Mismatches 10; Indels 0; Gaps 0;

Oy 1 TLWGFEVDGRSARGCOTSDAAAEALNASKEEALQIIRKIPKYLFOFHNLNSAI 60  
 Db 52 TLWGFEVDGRSARGCOTSDAAAEALNASKEEALQIIRKIPKYLFOFHNLNSAI 111  
 Oy 61 AAFDKTP 70  
 Db 112 DRIFKDP 121

RESULT 8  
 ID Q9YL44 PRELIMINARY; PRT; 233 AA.  
 AC Q9YL44;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
 DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).  
 GN ACL.  
 OS Macropitium golden mosaic geminivirus.  
 OC VIRUS: SMDNA viruses, Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=51676;  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-JAMAICA STRAIN 1;  
 RA Roye M.E.;  
 RL "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses  
 from Jamaica".  
 RN Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-JAMAICA STRAIN 1;  
 RA Roye M.E., McLaughlin W.A., Maxwell D.P.;  
 RT "Molecular characterization of two distinct geminiviruses infecting M.  
 acyroides from Jamaica";  
 RT Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.  
 RC EMBL: AF098940; A017850.1; -  
 DR INTERPRO: IP0001191; -  
 DR PFAM: PF00759; Gemini\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATL1.  
 FT NON\_TER 233  
 FT NON\_TER 233  
 SQ SEQUENCE 233 AA; 26355 MW; AA09AFA4D2166A02 CRC64;

Query Match 72.18; Score 259; DB 12; Length 233;  
 Best Local Similarity 68.68; Pred. No. 1e-21;



Viruses: ssDNA viruses: Geminiviridae; Begomovirus.

NCBI\_TaxID=67762;

SEQUENCE FROM N.A.

STRAIN=GORAL

Ooi K., Ohshita S., Ishii I., Yahara T.;

"Molecular phylogeny of geminivirus infecting wild plants in Japan.";

J. Plant Res. 110:247-257(1997).

J. EMBL: AB01303; BAA34010.1; -.

01-AUG-1998 (T-EMBLrel. 07, Created)

01-AUG-1998 (T-EMBLrel. 07, Last sequence update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

01-JUN-2000 (T-EMBLrel. 14, Last annotation update)

Query Match 69.4%; Score 249; DB 12; Length 354;  
Best Local Similarity 4.2%; Pred. No. 2.3e-20;  
Matches 49; Conservative 6; Mismatches 11; Indels 0; Gaps 0;  
QY 4 MGEFQVDSRGSGCOTNDAAAEALNASSKEALQITIREKIPKYLFOPHNLNSALAAI 63  
||||| 109 MGEFQVDSRGSGCOTNDAAAEALNASSKEALQITIREKIPKYLFOPHNLNSALAAI 168  
QY 64 FKTYPE 69  
Db 169 FURPFD 174

Search completed: February 3, 2001, 02:22:52  
Job time: 1869 sec

Result	Query No.	Score	Match	Query Length	ID	Description
1	342	95.3	352	1	VALL_TGMV	P03567 tomato gold
2	242	67.4	368	1	VALL_PYRV	P27258 potato yell
3	229	63.8	351	1	VALL_CIVK	P14982 cassava lat
4	229	63.8	358	1	VALL_CIVN	P14972 cassava lat
5	229	63.8	358	1	VALL_CIVK	P14972 cassava lat
6	216	60.2	345	1	VALL_BHCA	P08529 tomato yell
7	215	59.9	359	1	VALL_TYLCU	P38609 tomato yell
8	212	59.1	358	1	VALL_BCTV	P14991 beet curly
9	211	58.8	353	1	VALL_BGNV	P05175 bean golden
10	211	58.8	359	1	VALL_TYLCM	P27260 tomato yell
11	210	58.5	355	1	VALL_ABRWV	P21947 abutilon mo
12	207	57.7	361	1	VALL_TMOV	O06557 tomato yell
13	207	57.7	361	1	VALL_TMOV	O06557 tomato yell
14	107	29.8	347	1	VALL_SLSV	P29348
15	63.5	17.7	1610	1	CCAD_MESAV	O99244 mesocricetu
16	63.5	17.7	2161	1	CCAD_HAPIV	O101668 homo sapien
17	63.5	17.7	2203	1	CCAD_RAT	P277332 rattus norv
18	61	17.0	1483	1	CYPL_YEAST	P12351 saccharomy
19	59.5	16.6	224	1	SP2R_JACUS	P391551 bacillus su
20	59.5	16.6	4466	1	DYHC_TRIGR	P23098 tripneustes
21	59.5	16.6	4466	1	DYHC_TRIGR	P23098 tripneustes
22	59	16.4	387	1	KYPL_HBHSN	P35522 human papill
23	58.5	16.3	136	1	K452_CAREL	P025250 allobaccha
24	58.5	16.3	299	1	V175_HELPJ	O92mg7 halobacte
25	57.5	16.0	367	1	LHX4_MOUSE	P537776 mus musculu
26	57.5	16.0	1852	1	CCAS_CYP2A	P23216 cyprinus ca
27	57	15.9	247	1	YCPA_YEAST	P25349 saccharomy
28	57	15.9	256	1	YKGP_YEAST	P34247 saccharomy
29	56.5	15.7	2180	1	CCAD_CHICK	O77300 gallus gall
30	56.5	15.7	2180	1	CCAD_CHICK	O77300 gallus gall
31	56	15.6	295	1	V175_HESAV	P186971 melitococc
32	56	15.6	295	1	V175_HESAV	P186971 melitococc
33	56	15.6	295	1	VALL_TYDVA	P31617 homo sapien
34	56	15.6	836	1	GC5R_HUMAN	O99052 homo sapien

AC P27258:  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DE AL1 PROTEIN (Rel. 23, Last annotation update)  
 OS Potato yellow mosaic virus (Isolate Venezuela).  
 OS Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-91311403; PubMed-1856690;  
 RA Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.:  
 RT "The nucleotide sequence of the infectious cloned DNA components of  
 RT potato yellow mosaic virus." J. Gen. Virol. 72:1515-1520(1991).  
 RL J. Gen. Virol. 72:1515-1520(1991).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: D00940; BAA00782.1; -  
 DR PIR: J00364; Q0CVPT.  
 DR INTERPRO: IPR001191; -  
 DR PRAM: P00759; GeminA1; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR PRINTS: PR00228; GEMCOATCLVL1.  
 KW ATP-binding.  
 KW NP\_BIND 222 229 ATP (POTENTIAL).  
 FT SEQUENCE 361 AA; 40850 MW; 5627A33BF1264383 CRC64;  
 SQ

Query Match 57.4%; Score 242; DB 1; Length 361;  
 Best Local Similarity 66.7%; Pred. No. 3c-20;  
 Matches 46; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 1 TLVWGFEQVDSRGSGCOTSDNDAAALNASKEALQIREKIPKIFYLFQHNLSAL 60  
 DB 110 TVWGFEQVDSRGSGCOTSDNDAAALNASKEALQIREKIPKIFYLFQHNLSNL 169  
 QY 61 AALFKPTKE 69  
 DB 170 DRIFNKAPE 178

RESULT 3  
 AC VAL1 CLVKN STANDARD; PRT: 358 AA.  
 AC P14582;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE AL1 PROTEIN (40.4 KDA PROTEIN).  
 OS Cassava latent virus (strain West Kanyap 84).  
 OS Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX Stanley J., Gay M.R.:  
 RA "Nucleotide sequence of cassava latent virus DNA.\*";  
 RL Nature 301:260-262(1983).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: J02057; -; NOT\_ANNOTATED\_CDS.  
 DR INTERPRO: IPR001191; -  
 DR PRAM: P00759; GeminA1; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR PRINTS: PR00228; GEMCOATCLVL1.  
 KW ATP-binding.  
 KW NP\_BIND 220 227 ATP (POTENTIAL).  
 FT SEQUENCE 358 AA; 40346 MW; ED173E75FE5D49 CRC64;  
 SQ

Query Match 63.8%; Score 229; DB 1; Length 358;  
 Best Local Similarity 60.0%; Pred. No. 8.9e-19;  
 Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

QY 1 TLVWGFEQVDSRGSGCOTSDNDAAALNASKEALQIREKIPKIFYLFQHNLSAL 60  
 DB 109 TVWGFEQVDSRGSGCOTSDNDAAALNASKEALQIREKIPKIFYLFQHNLSNL 168  
 QY 61 AALFKPTKEP 70  
 DB 169 DRIFQEPAP 178

RESULT 4  
 AC VAL1 CLVKN STANDARD; PRT: 358 AA.  
 AC P14972;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE AL1 PROTEIN (40.4 KDA PROTEIN).  
 OS Cassava latent virus (strain Nigerian).  
 OS Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-90174930; PubMed-2308831;  
 RA Morris B., Coates L., Lowe S., Richardson K., Eddy P.:  
 RT "Nucleotide sequence of the infectious cloned DNA components of  
 RT African cassava mosaic virus (strain Nigerian).";  
 RL Nature 341:197-199(1990).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL: X17095; CAA34953.1; -  
 DR PIR: S07594; S07594.  
 DR INTERPRO: IPR001191; -  
 DR PRAM: P00759; GeminA1; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR PRINTS: PR00228; GEMCOATCLVL1.  
 KW ATP-binding.  
 KW NP\_BIND 220 227 ATP (POTENTIAL).  
 FT SEQUENCE 358 AA; 40435 MW; 1DB16B80CB2D5E2C CRC64;  
 SQ

Query Match 63.8%; Score 229; DB 1; Length 358;  
 Best Local Similarity 60.0%; Pred. No. 8.9e-19;  
 Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

QY 1 TLVWGFEQVDSRGSGCOTSDNDAAALNASKEALQIREKIPKIFYLFQHNLSAL 60  
 DB 109 TVWGFEQVDSRGSGCOTSDNDAAALNASKEALQIREKIPKIFYLFQHNLSNL 168



```

QY 61 AAIFDKTPEP 70
Db 169 DRIFQEPAP 178

RESULT 5
VALL TYLCA
ID VALL TYLCA STANDARD: PRT; 362 AA.
AC P36279; 1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN (C1 PROTEIN).
CN C1.
OS Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RX MEDLINE=93139776; PubMed=8473446;
RA DRY I.B., Bidden J.F., Krake L.R., Mullineux P.M., Rezaian M.A.;
RT "Nucleotide sequence and genome organization of tomato leaf curl
geminiavirus."
RL J. Gen. Virol. 74:147-151(1993).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC EMBL: X70418; CAA49856.1;
DR PIR: S31875; S31875.
DR JQ2300; JQ2300.
DR INTERPRO: IPR001191;
DR PFAM: PF00799; Gemini_A1; 1.
DR PRINTS: PR00227; GEMCOATALL.
DR PRINTS: PR00228; GEMCOATCLV1.
KW ATP-binding.
FT NP-BIND 221 228 ATP (BY SIMILARITY).
SQ SEQUENCE 362 AA: 41197 MW: 34367184B4704098 CRC64:
Query Match 61.8%; Score 222; DB 1; Length 362;
Best Local Similarity 52.9%; Pred. No. 5.6e-18;
Matches 4%; Conservative 9; Mismatches 15; Indels 16; Gaps 1;
QY 1 TLVNGSFQVQDSARGCQTSNDAAALNASSKEALQIIRKPEKLYQFPHNLSNL 60
Db 110 TLVNGSFQVQDSARGCQTSNDAAALNASSKEALQIIRKPEKLYQFPHNLSNL 169
QY 61 AAIFDKTPEP 70
Db 170 DRIFQEPAP 178

RESULT 7
VALL TYLCA
ID VALL TYLCA STANDARD: PRT; 359 AA.
AC P38609; 1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE AL1 PROTEIN (C1 PROTEIN).
CN C1.
OS Tomato yellow leaf curl virus (strain Murcia) (TYLCV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RX MEDLINE=94256836; PubMed=8198442;
RA Norris E., Hidalgo E., Accotto G., Moriones E.;
RT "High similarity among the tomato yellow leaf curl virus isolates
from the west Mediterranean basin: the nucleotide sequence of an
Arabidopsis chloroplast virus."
PL Arch. Virol. 135:165-170(1994).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC EMBL: 225751; CAA81026.1;
DR PIR: S39211; S39211.
DR INTERPRO: IPR001191;

```

```

DR INTERPRO: IPR001301;
DR PFAM: PF00759; Gemini_ALI; 1.
DR PRINTS: PR00227; GEMCOATLVL.
DR AT: 228 ATP (POTENTIAL).
DR NP_BIND 222 228 ATP (POTENTIAL).
DR SEQUENCE 359 AA; 41065 MW; 20170A51BF50A3BC CRC64;

Query Match
Best Local Similarity 59.9%; Score 215; DB 1; Length 359;
Matches 40; Conservative 9; Mismatches 20; Indels 0; Gaps 0;

QY 2 LVWGFQVDGRSARGGQTSNDAAAEALNASKEEALQITREKIPKYLFOFHINLSAL 61
DB 111 LVWGFQVDGRSARGGQTSNDAAAEALNASKEEALQITREKIPKYLFOFHINLSAL 170
QY 62 AAFDKTPEP 70
DB 171 RVQFFPAP 179

RESULT 8
VALI_BCTV
ID VALI_BCTV STANDARD: PRT; 358 AA.
AC PI4591;
DT 01-APR-1990 (Rel. 14, Created)
DR PRINTS: PR00227; GEMCOATLVL.
DE ALL PROTEIN (40.8 KDA PROTEIN).
OS Beet curly top virus (BCTV).
OC Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
RN
RP SEQUENCE FROM N.A.
RA Stanley J., Markham P.G., Callis R.J., Pinner M.S.;
RT best curly top virus.
RL EMBL J. 5:1761-1767(1986).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is not modified and
CC this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; X04144; -; NOT_ANNOTATED_CDS.
DR INTERPRO: IPR001191;
DR PFAM: PF00759; Gemini_ALI; 1.
DR PRINTS: PR00227; GEMCOATLVL.
DR AT: 222 229 ATP (POTENTIAL).
DR NP_BIND 222 229 ATP (POTENTIAL).
DR SEQUENCE 358 AA; 40889 MW; 39A45FE3C0B9C333 CRC64;

Query Match
Best Local Similarity 59.1%; Score 212; DB 1; Length 358;
Matches 39; Conservative 15; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLWGFQVDGRSARGGQTSNDAAAEALNASKEEALQITREKIPKYLFOFHINLSAL 60
DB 110 TLWGFQVDGRSARGGQTSNDAAAEALNASKEEALQITREKIPKYLFOFHINLSAL 169
QY 61 AAFDKTPEP 70
DB 170 QKIFORPPDP 179

RESULT 9
VALI_BGMV
ID VALI_BGMV STANDARD: PRT; 353 AA.
AC POS175;
DT 13-AUG-1987 (Rel. 05, Created)
DR PRINTS: PR00227; GEMCOATLVL.
DE ALL PROTEIN (40.2 KDA PROTEIN).
OS Bean golden mosaic virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN
RP SEQUENCE FROM N.A.
RA J. Bossert M., Goodman R.M.;
RT fluctuation in geminivirus and a model for gene
RT Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is not modified and
CC this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; M10070; AAA46318.1;
DR INTERPRO: IPR001301;
DR PFAM: PF00759; Gemini_ALI; 1.
DR PRINTS: PR00227; GEMCOATLVL.
DR AT: 222 229 ATP (POTENTIAL).
DR NP_BIND 222 229 ATP (POTENTIAL).
DR SEQUENCE 353 AA; 40190 MW; 80FA79DF6029A34 CRC64;

Query Match
Best Local Similarity 58.8%; Score 211; DB 1; Length 353;
Matches 41; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLWGFQVDGRSARGGQTSNDAAAEALNASKEEALQITREKIPKYLFOFHINLSAL 60
DB 110 TLWGFQVDGRSARGGQTSNDAAAEALNASKEEALQITREKIPKYLFOFHINLSAL 169
QY 61 AAFDKTPEP 70
DB 170 ERIFFKVPPEP 179

RESULT 10
VALI_TYLCM
ID VALI_TYLCM STANDARD: PRT; 359 AA.
AC P27260;
DT 01-AUG-1992 (Rel. 23, Created)
DR PRINTS: PR00227; GEMCOATLVL.
DE ALL PROTEIN (CI PROTEIN).
OS Tomato yellow leaf curl virus (strain Marmande) (TYLCV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN
RP SEQUENCE FROM N.A.
RA MEDLINE-92107660; PubMed-1840676;
RA Kheyr-Pour A., Bendahmane M., Matzelt V., Accotto G.P., Crespi S.,
RA von Borstel B., leaf curl virus from Sardinia is a
RA whitefly-transmitted monopartite geminivirus.
RT Nucleic Acids Res. 19:6763-6769(1991).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/annouce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).)

CC EMBL: X61153; CAA43466.1; -;  
 DR PIR: S22593; S22593.  
 DR INTERPRO: IPR001191; -;  
 DR INTERPRO: IPR001301; -;  
 DR PFAM: PF00799; Gemini\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATL1.  
 KW ATP-binding, 220 227 ATP (POTENTIAL).  
 FT NP\_BIND 220 227  
 SQ SEQUENCE 355 AA: 40733 MW: 971784A07C93EFA7 CRC64;

Query Match 58.8%; Score 211; DB 1; Length 359;  
 Best Local Similarity 58.8%; Pred No. 1.3e-16;  
 Matches 39; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

Oy 2 LWNKGFQVDRSARGCQTSNDAAAEALNASSKEALQIRKIPKYLQFQHNLSAL 61  
 Db 111 LWNKGFQVDRSARGCQTSNDAAAEALNASSKEALQIRKIPKYLQFQHNLSAL 61  
 Oy 62 AIFDKTPP 70  
 Db 171 KVFQVPAP 179

RESULT 11  
 VALL\_ABMVW STANDARD: PRT; 355 AA.

AC P21947; 991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE ALL PROTEIN.  
 GN Abutilon mosaic virus (isolate West India).  
 OS Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 RP MEDLINE-91020984; PubMed-2219703;  
 RA Frischmuth T., Zimmat G., Jeske H.;  
 RT "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic  
 RL virology 178:461-468(1990).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/annouce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).)

CC EMBL: X15983; -; NOT\_ANNOTATED\_CDS.  
 DR PIR: A36214; OCQVM1.  
 DR INTERPRO: IPR001191; -;  
 DR INTERPRO: IPR001301; -;  
 DR PFAM: PF00799; Gemini\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATL1.  
 KW ATP-binding, 221 228 ATP (POTENTIAL).  
 FT NP\_BIND 221 228  
 SQ SEQUENCE 355 AA: 40257 MW: 16A2CABA63251B95 CRC64;

Query Match 58.5%; Score 210; DB 1; Length 355;  
 Best Local Similarity 57.1%; Pred. No. 1.3e-16;

Matches 40; Conservative 11; Mismatches 19; Indels 0; Gaps 0;  
 Oy 1 TLVNGFQVDRSARGCQTSNDAAAEALNASSKEALQIRKIPKYLQFQHNLSAL 60  
 Db 110 TLVNGFQVDRSARGCQTSNDAAAEALNASSKEALQIRKIPKYLQFQHNLSAL 60

Oy 61 AAIDKTPP 70  
 Db 170 ERIFAKAP 179

RESULT 12  
 VALL\_JMOV STANDARD: PRT; 361 AA.

AC O06657;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE ALL PROTEIN.  
 GN Tomato mottle virus (isolate Florida) (TMoV).  
 OS Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 RP MEDLINE-93107858; PubMed-1469361;  
 RA Abouzid A.M., Polston J.E., Hiebert E.;  
 RT "The nucleotide sequence of tomato mottle virus, a new geminivirus  
 RT isolated from tomatoes in Florida."  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/annouce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).)

CC EMBL: L14460; AAC32414.1; -;  
 DR PIR: JQ1870; JQ1870.  
 DR INTERPRO: IPR001191; -;  
 DR INTERPRO: IPR001301; -;  
 DR PFAM: PF00799; Gemini\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATL1.  
 KW ATP-binding, 222 229 ATP (BY SIMILARITY).  
 FT NP\_BIND 222 229  
 SQ SEQUENCE 361 AA: 40516 MW: 8138665CEBAC6950 CRC64;

Query Match 57.7%; Score 207; DB 1; Length 361;  
 Best Local Similarity 54.3%; Pred No. 2.8e-16;  
 Matches 38; Conservative 15; Mismatches 17; Indels 0; Gaps 0;

Oy 1 TLVNGFQVDRSARGCQTSNDAAAEALNASSKEALQIRKIPKYLQFQHNLSAL 60  
 Db 110 TLVNGFQVDRSARGCQTSNDAAAEALNASSKEALQIRKIPKYLQFQHNLSAL 60

Oy 61 AAIDKTPP 70  
 Db 170 ERIFAKAP 179

RESULT 13  
 VALL\_TYICV STANDARD: PRT; 357 AA.

AC P27259;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE ALL PROTEIN (C1 PROTEIN).  
 GN C1.



Query Match	Best Local Similarity	7.74%; DB 1; Length 1610;
Matches 21; Conservative	5; Mismatches 19; Indels 29; Gaps	
706 WNAWVDGIMVGGSSSSGICVYPIILPGNVI LLWFLAIVDNLADESLNTAQK 765	4 WEEFQVDSRSGAGQGS-----	-----NDAAEALNASK 34
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

Search completed: February 3, 2001, 02:24:28  
Job time: 640 sec

---

**THIS PAGE BLANK (USPTO)**



```

Query Match      67.4%: Score 242; DB 1; Length 361;
Best Local Similarity 66.74; Pred. No. 5, 5e-20;
Matches 46; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

OY 1 TLVNGEFOVGRSARGGCGTSDNAAALNASSKEALQIIREKIPKYLFOFHNLSAL 60
   1 : |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 110 TLEWGEFOIDGRSARGGCGTVDNAAALNASSKEALQIIREKLPKFLFOVHLSNCL 169
   1 : |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 61 AAIIDKTPPE 59
Db 170 DRIFQPPAP 178

RESULT 3
hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)
C:Species: Cassava latent virus
C:Date: 17-Feb-1994 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C:Accession: S07594
R:Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.
Nucleic Acids Res. 18, 197-198, 1990
A:Title: Nucleotide sequence of the infectious cloned DNA components of African cassava
A:Reference number: S07594
A:Accession: S07594
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-358 <NOR>
A:Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376
C:Genetics:
A:Map position: segment DNAL
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match      63.8%: Score 229; DB 2; Length 358;
Best Local Similarity 60.04; Pred. No. 1, 6e-18;
Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

OY 1 TLVNGEFOVGRSARGGCGTSDNAAALNASSKEALQIIREKIPKYLFOFHNLSAL 60
   1 : |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 109 TVNGEFOIDGRSARGGCGSANDATYAKALNSASAEALQIIREKIPKYLFOFHNLSAL 166
   1 : |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 61 AAIIDKTPPE 70
Db 169 DRIFQPPAP 178

RESULT 4
hypothetical protein - tomato yellow leaf curl virus (strain Australia)
N:Alternate names: CL protein
C:Species: tomato yellow leaf curl virus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C:Accession: J01887
R:Dry, I.B.; Mogen, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezanian, M.A.
J. Gen. Virol. 74, 147-151, 1993
A:Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.
A:Reference number: J01887
A:Accession: J01887
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-362 <DR>
A:Cross-references: GB:S53251
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match      61.8%: Score 222; DB 1; Length 362;
Best Local Similarity 52.94; Pred. No. 1e-17;
Matches 45; Conservative 9; Mismatches 15; Indels 16; Gaps 1;

OY 1 TLVNGEFOVGRSARGGCGTSDNAAALNASSKEALQIIREKIPKYLFOFHNLSAL 60
   1 : |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db 110 TLEWGEFOIDGRSARGGCGSANDATYAKALNTGKSEALNVLRELAPKQVLOFHNLSNL 169
OY 61 AAI-----FDKTPPE 69
   1 : |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 170 DRIFQPPAPVYVPSFLSSFDKTPPE 194
   1 : |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 5
replicase - pepper huasteco virus (component A)
N:Alternate names: ORF AL1 protein
C:Species: pepper huasteco virus
C:Date: 22-Nov-1995 #sequence_revision 14-Jul-1994 #text_change 20-Sep-1999
C:Accession: J02300
R:Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante,
J. Gen. Virol. 74, 2225-2231, 1993
A:Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b
A:Reference number: J02299; MUID:94015007
A:Accession: J02300
A:Molecule type: DNA
A:Residues: 1-900
A:Cross-references: GB:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match      60.2%: Score 216; DB 2; Length 349;
Best Local Similarity 58.6%: Pred. No. 4, 8e-17;
Matches 41; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

OY 1 TLVNGEFOVGRSARGGCGTSDNAAALNASSKEALQIIREKIPKYLFOFHNLSAL 60
   1 : |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 110 TVNGEFOIDGRSARGGCGSANDTAYAKALNSASAEALQIIREKIPKYLFOFHNIVNSA 169
   1 : |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 61 AAIIDKTPPE 70
Db 170 NRIFQTPPE 179

RESULT 6
AL1 protein - pepper rizado amarillo virus
C:Species: pepper rizado amarillo virus
C:Date: 22-Nov-1995 #sequence_revision 26-May-1995 #text_change 20-Sep-1999
C:Accession: S31875
R:Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Jimenez, B.; Herrera-Estrella, L.; Rivera
submitted to the EMBL Data Library, February 1993
A:Description: Complete nucleotide sequence of pepper huasteco virus: analysis and co
A:Reference number: S31872
A:Accession: S31875
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-349
A:Cross-references: EMBL:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match      60.2%: Score 216; DB 2; Length 349;
Best Local Similarity 58.6%: Pred. No. 4, 8e-17;
Matches 41; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

OY 1 TLVNGEFOVGRSARGGCGTSDNAAALNASSKEALQIIREKIPKYLFOFHNLSAL 60
   1 : |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 110 TVNGEFOIDGRSARGGCGSANDTAYAKALNSASAEALQIIREKIPKYLFOFHNIVNSA 169
   1 : |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 61 AAIIDKTPPE 70
Db 170 NRIFQTPPE 179

RESULT 7
gene C1 protein - tomato yellow leaf curl virus

```



C:Species: tomato yellow leaf curl virus  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 20-Sep-1999  
 C:Accession: S39211  
 R:Abouzeid, S.; Hladik, B.; Accotto, G.; Moriones, E.  
 A:Description: High similarity among the tomato yellow leaf curl virus isolates from the Mediterranean area.  
 A:Reference number: S39209  
 A:Accession: S39211  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-359 <NOR>  
 A:Cross-references: EMBL:Z25751; NID:g433655; PIDN:CA81026.1; PID:g433658  
 C:Superfamily: tomato golden mosaic virus A11 protein

Query Match 59.9%; Score 215; DB 2; Length 359;  
 Best Local Similarity 58.0%; Pred. No. 6.5e-17;  
 Matches 40; Conservative 9; Mismatches 20; Indels 0; Gaps 0;

QY 2 TLVNGEFQVGRSARGGCGTSDAAAEALNASSKEEALQITREKIPKYLQFPHNLSAL 61  
 DB 111 LKMGTFQIDGSRAGGCGTQANDAYAKALNAGSKSEALDVIKELAPROYLILHFHINSNLD 170

QY 62 AIFDKTPEP 70

DB 171 RVFQVPPAP 179

# RESULT 8

QY A11 protein - beet curly top virus  
 C:Species: beet curly top virus  
 C:Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 20-Sep-1999  
 C:Accession: S28360  
 R:Scantney, J.; Markham, B.G.; Callis, R.J.; Pinner, M.S.  
 A:Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly top virus.  
 A:Reference number: S28360  
 A:Accession: S28360  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-385 <STA>  
 A:Cross-references: GB:FM4597; EMBL:X04144; NID:g210678; PIDN:AAA2751.1; PID:g210679  
 C:Superfamily: tomato golden mosaic virus A11 protein

Query Match 59.1%; Score 212; DB 2; Length 385;  
 Best Local Similarity 55.7%; Pred. No. 1.5e-16;  
 Matches 39; Conservative 15; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLVNGEFQVGRSARGGCGTSDAAAEALNASSKEEALQITREKIPKYLQFPHNLSAL 60  
 DB 137 TLNGEFQIDGSRAGGCGTQANDSYAKALNLSLQALQILKEQPKYFQVHLHNLNNA 196

QY 61 AAIKFKTPEP 70

DB 197 QRIFQRPDP 206

# RESULT 9

QY hypothetical protein C4 - tomato yellow leaf curl virus  
 C:Species: tomato yellow leaf curl virus  
 C:Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 20-Sep-1999  
 C:Accession: S22593  
 R:Kheyr-Pour, A.; Bendahmane, M.; Matzeit, V.; Accotto, G.P.; Crespi, S.; Gronenborn, B.  
 A:Title: The nucleotide sequence of the tomato yellow leaf curl virus (TYLCV) from Sardinia is a whitefly-transmitted monopartite virus.  
 A:Reference number: S22598; NID:92107660  
 A:Accession: S22593  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-359 <KHE>

A:Cross-references: EMBL:X61153; NID:g62211; PIDN:CAA43466.1; PID:g62217  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991  
 C:Superfamily: tomato golden mosaic virus A11 protein

Query Match 58.8%; Score 211; DB 2; Length 359;  
 Best Local Similarity 56.5%; Pred. No. 1.9e-16;  
 Matches 39; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

QY 2 TLVNGEFQVGRSARGGCGTSDAAAEALNASSKEEALQITREKIPKYLQFPHNLSAL 61  
 DB 111 LKMGTFQIDGSRAGGCGTQANDAYAKALNAGSKSEALDVIKELAPROYLILHFHINSNLD 170

QY 62 AIFDKTPEP 70

DB 171 KVFQVPPAP 179

# RESULT 10

QY A11 protein - abutilon mosaic virus (isolate West India)  
 C:Species: abutilon mosaic virus  
 C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 08-Apr-1994  
 C:Accession: A36214  
 R:Frischmuth, T.; Zimmat, G.; Jeske, H.  
 A:Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as well as eukaryotic features.  
 A:Reference number: A36214; NID:91020984  
 A:Accession: A36214  
 A:Molecule type: DNA  
 A:Residues: 1-355 <PRI>  
 A:Cross-references: EMBL:X15983  
 C:Genetics:  
 A:Map position: segment A  
 C:Superfamily: tomato golden mosaic virus A11 protein

Query Match 58.5%; Score 210; DB 1; Length 355;  
 Best Local Similarity 57.1%; Pred. No. 2.4e-16;  
 Matches 40; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

QY 1 TLVNGEFQVGRSARGGCGTSDAAAEALNASSKEEALQITREKIPKYLQFPHNLSAL 60  
 DB 110 TANGFQIDGSRAGGCGTQANDSYAKALNAGDVQSLNHLKKEQPKYQVQVQNNIRNSL 169

QY 61 AAIKFKTPEP 70

DB 170 ERIFAKAPEP 179

# RESULT 11

QY A11 protein - tomato mottle virus (isolate Florida)  
 C:Species: tomato mottle virus  
 C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 07-May-1999  
 C:Accession: J01870  
 R:Abouzeid, A.M.; Polston, J.E.; Hiebert, E.  
 A:Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated from tomato.  
 A:Reference number: J01870; NID:93107858  
 A:Accession: J01870  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-358 <ABO>  
 A:Cross-references: GB:L14460  
 C:Genetics:  
 A:Map position: segment A  
 C:Superfamily: tomato golden mosaic virus A11 protein

Query Match 57.7%; Score 207; DB 1; Length 358;  
 Best Local Similarity 54.3%; Pred. No. 5.3e-16;  
 Matches 38; Conservative 15; Mismatches 17; Indels 0; Gaps 0;

OY	1	TLVCEKPOVCHSRANGCOTNDAAAPALAAASKEALOILIEKIPKYLQPHNLSAL	50
DB	107	TEIWDQPIIDGRSANGQSNDYSAKALAAQSQSLAVIREEOPKPYLQNHITSNL	166
OY	61	AATFDKTPRP	70
DB	167	ERENAKPER	176

RESULT 12

JQ2327

All protein - Indian cassava mosaic virus

Alt:Alternate names: replication-associated protein

C:Species: Indian cassava mosaic virus

C:Date: 28-Aug-1985 sequence\_revision 07-Oct-1994 text\_change 20-Sep-1999

Accession: F05883; S5583; J

RefSeq: F05883; Robinson, B. D.

J. Gen. Virol. 74, 2437-2443, 1993

A:Title: Nucleotide sequence evidence for the occurrence of three distinct w

A:Reference number: JQ2326; MUID:94065670

A:Accession: JQ2327

A:Molecule type: DNA

A:Accession: JQ2327

A:Accession: JQ2327

A:CrossReferences: EMBL: Z24758; NID: g395751; PIDN: CAAR0891.1; PID: g584046

A:Superfamily: Tomato golden mosaic virus All protein

Query Match 57.1%; Score 205; DB 2; Length 351;  
Best Local Similarity 59.7%; Pred. No. 8.7e-16;  
Matches 40; Conservative 9; Mismatches 18; Indels 0; Gaps 0;

[illegible]

13 RESULT  
S39235 gene C1 protein - tomato yellow leaf curl virus  
S39235 C-Species: tomato yellow leaf curl virus  
C-Date: 06-Jan-1995 frequency-revision 06-Jan-1995 #text\_change 08-Sep-1997  
R-Accession: S39235  
R-Accession: S39235 E: Valira, E.; Bosco, D.; Accotto, G.  
Submitted to the EMBL Data Library, December 1993  
A-Description: A cloned DNA from a TYLCV isolate from Sicily showing low in-  
A-Reference number: S39233  
A-Accession: S39233  
A-Status: preliminary  
A-Molecule type: DNA  
A-Feature key: CDS  
A-Feature key: ORF  
A-Superfamily: EMBL:228490; NTD:q1041671; PID:q1334964  
A-Superfamily: tomato yellow mosaic virus A11 protein

Query Match 57.1%; Score 205; DB 2; Length 359;  
Best Local Similarity 55.1%; Pred. No. 9e-16;  
Matches 38; Conservative 11; Mismatches 20; Indels 0; Gaps 0;

```

QY      2  LVN6F0V0GSRGARGCGTSDNMAAANASSKKEALQITREKTYLFOFHNSALA 61
DB      111  LKNGTFOIGSRGARGCGTNDAYAKANARKSKSEALDYKQLAPDYVJFHNTSSNLD 170

QY      62  AIFDKTEP 70
DB      171  KVFQFPAP 179

```

RESULT 14

S59885  
 replication-associated protein C1 - tomato yellow leaf curl virus  
 C:Species: tomato yellow leaf curl virus  
 C:SpecDate: 14-Jan-1996 /sequence\_revision 01-Mar-1996 /text\_change 20-Sep-1999  
 C:Accession: S59885  
 R:Hong, Y.; Harrison, B.D.  
 A:Description: Nucleotide sequences from tomato leaf curl viruses from different coun-  
 d geminiviruses.  
 A:Reference number: S58346  
 A:Accession: S59885  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Sequence length: 600 nt  
 A:C:References: EMBL:Z48182; NID:q944838; PTDN:CAA8229.1.; PID:q974211  
 C:Superfamily: Superfamily: tomato yellow mosaic virus AL1 protein

	Query Match	54.3%	Score 195	DB 2	Length 360
	Best Local Similarity	57.6%	Pred. No. 1,2e-14		
	Matches 38	Conservative 12	Mismatches 16	Indels 0	Gaps 0
QY	4	WEEVDGSRGAGCGQTSNDAAALNASKKPAQIIRKIPKRYLPQPHNLNAAAI	63		
DB	113	FGVVDGSRGAGCGQTSNDATAACINSGKAADILREKAPFVLPQPHNLNLDRI	172		
QY	64	FKTPEE 69			
DB	173	FYPSAE 178			

	RESULT	15
	QOCVC1	
A:	All protein - tomato yellow leaf curl virus	
N:	Alternate names: Cl protein	
C:	Species: tomato yellow leaf curl virus	
D:	Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999	
C:	Accession: D40779	
R:	Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.	
V:	Virology 185, 151-161, 1991	
T:	Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin	
R:	Reference number: A40779; MID:52024070	
A:	Accession: D40779	
S:	Status: translation not shown	
P:	Protein Data Bank	
R:	Release date: 11-357-NAV>	
C:	Cross-references: GB:I5656; PID:g62204; PIDN:CAA33688.1; PID:g62207	
F:	Family: tomato golden mosaic virus All protein	

```

Query Match      53.8%  Score 193;  DB 1;  Length 357;
Best Local Similarity 62.3%  Pred. No. 2.le-14;
Matches 38;  Conservative 10;  Mismatches 13;  Indels 0;  Gaps 0;

Qy  4  WGFVQVDSRGARGCOTSDAAALAAALMASKEEAALQIREKIPKYLPOFHNLSALAAI  63
      :  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db  111  FGVSQIDGSRGARGQQSDAAAYALNSGSKSALNILEAKPQDYLIQPHNLSNDIRI  170

```

Search completed: February 3, 2001, 02:17:41  
Job time: 2213 sec

GenCorte version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:19:47 ; Search time 109.09 Seconds  
(without alignments)  
11,523 Million cell updates/sec

Title: US-09-289-346A-7

Perfect score: 359

Sequence: 1 TLVNGFEQVDSRGSGCQT.....PQFINLSALAAIFDXTPET 70

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957043 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents AA\*

1: /cgn2\_5/ptodata/1/aa/5A\_COMB.pep.\*

2: /cgn2\_5/ptodata/1/aa/5B\_COMB.pep.\*

3: /cgn2\_5/ptodata/1/aa/6\_COMB.pep.\*

4: /cgn2\_5/ptodata/1/aa/PCTUS\_COMB.pep.\*

5: /cgn2\_5/ptodata/1/aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	211	58.8	359	3	US-08-809-103B-2
2	211	58.8	359	3	US-08-809-103B-4
3	211	58.8	359	3	US-08-809-103B-6
4	211	58.8	359	3	US-08-809-103B-8
5	63.5	17.7	2161	1	US-07-745-206A-2
6	63.5	17.7	2161	1	US-08-455-543A-29
7	63.5	17.7	2161	1	US-08-455-543A-51
8	63.5	17.7	2161	1	US-08-223-305C-49
9	63.5	17.7	2161	2	US-08-223-305C-51
10	63.5	17.7	2161	2	US-08-311-363-2
11	58	16.2	374	2	US-08-928-692-51
12	57	15.9	454	3	US-08-446-100-26
13	57	15.9	454	3	US-08-446-100-27
14	57	15.9	454	3	US-08-446-100-28
15	57	15.9	454	3	US-08-446-100-29
16	57	15.9	454	3	US-08-446-100-30
17	56	15.6	602	2	US-08-415-652-6
18	56	15.6	771	1	US-07-923-976-6
19	56	15.6	783	5	5422248-2
20	56	15.6	836	1	US-07-923-976-4
21	56	15.6	863	1	US-07-923-976-9
22	53.5	15.5	863	1	US-08-468-924-3
23	53.5	15.5	860	1	US-08-468-924-3
24	55	15.3	432	1	US-08-165-919A-11
25	55	15.3	432	3	US-08-715-106-11
26	55	15.3	454	3	US-08-446-100-31
27	54.5	15.2	1713	3	US-08-600-982-24
28	54.5	15.2	1713	4	PCT-US94-10261A-24

29 54 15.0 751 3 US-09-036-987A-24  
30 53.5 14.9 373 3 US-08-258-287B-42  
31 53.5 14.9 373 3 US-08-368-704C-42  
32 53.5 14.9 3898 2 US-08-876-991-2  
33 53 14.8 3898 2 US-08-876-991-2  
34 53 14.8 284 2 US-08-423-641B-5  
35 53 14.8 284 2 US-08-820-980-5  
36 53 14.8 284 2 US-08-825-439-5  
37 53 14.8 824 1 US-08-221-750A-3  
38 52 14.6 480 1 US-07-803-636A-2  
39 52 14.5 244 1 US-08-910-973-22  
40 52 14.5 482 1 US-07-782-881A-1  
41 52 14.5 482 1 US-07-782-881A-1  
42 52 14.5 482 1 US-08-142-439A-7  
43 52 14.5 482 2 US-08-869-477-7  
44 52 14.5 1012 2 US-08-680-326-34  
45 51 14.2 237 1 US-08-910-973-13

#### ALIGNMENTS

RESULT 1  
US-08-809-103B-2  
: Sequence 2, Application US/0809103B  
: Patent No. 6133505  
: GENERAL INFORMATION:  
: INVENTOR: THOMPSON, BRUNO  
: TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT  
: NUMBER OF SEQUENCES: 17  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: YOUNG & THOMPSON  
: STREET: 745 South 23rd Street  
: CITY: Arlington  
: STATE: Virginia  
: COUNTRY: U.S.A.  
: ZIP: 22202  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.30  
: COUNTRY OF ORIGIN: U.S.A.  
: APPLICATION NUMBER: US/08/809,103B  
: FILING DATE: 17-MAR-1997  
: CLASSIFICATION: 800  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: FR 94.11040  
: FILING DATE: 15-SEP-1994  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: NO PCT/FR95/01192  
: FILING DATE: 15-SEP-1995  
: ATTORNEY/AGENT INFORMATION:  
: NAME: PATCH, Andrew J.  
: REGISTRATION NUMBER: 32,925  
: REFERENCE/DOCKET NUMBER: US8944L CNR TOM  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (703) 685-24297  
: TELEFAX: (703) 685-24297  
: TELEX: 248425 EPMON  
: INFORMATION FOR SEO ID NO: 2:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 359 amino acids  
: TYPE: amino acid  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
US-08-809-103B-2

Query Match 58.8%; Score 211; DB 3; Length 359;  
Best Local Similarity 56.5%; Pred. No. 2.6e-20;  
.Matches 39; Conservative 10; Mismatches 20; Indels 0; Gaps 0;



```

1  COUNTRY: U.S.A.
2  ZIP: 22202
3  COMPUTER READABLE FORM:
4  MEDIUM TYPE: Floppy disk
5  OPERATING SYSTEM: IBM compatible
6  SOFTWARE: PC-DOS/MS-DOS
7  CURRENT APPLICATION INFORMATION:
8  APPLICATION NUMBER: US/08/809,103B
9  FILING DATE: 17-MAR-1997
10 CLASSIFICATION: 800
11 PRIOR APPLICATION DATA:
12 APPLICATION NUMBER: FR 94,11040
13 FILING DATE: 15-SEP-1994
14 PRIOR APPLICATION NUMBER: WO PCT/FR95/01192
15 FILING DATE: 15-SEP-1995
16 ATTORNEY/AGENT INFORMATION:
17 NAME: PATCH, Andrew J.
18 REFERENCE/DOCKET NUMBER: US894AL CNR TOM
19 TELECOMMUNICATION INFORMATION:
20 TELEPHONE: (703) 521-2297
21 TELEFAX: (703) 685-0573
22 TELEX: 248425 EMBON
23 INFORMATION FOR SEQ ID NO: 8:
24 SEQUENCE CHARACTERISTICS:
25 LENGTH: 2161 amino acids
26 TYPE: amino acid
27 TOPOLOGY: linear
28 MOLECULE TYPE: protein
29 US-08-809-103B-8

Query Match 58.8%; Score 211; DB 3; Length 359;
Best Local Similarity 56.5%; Pred. No. 2,6e-70;
Matches 39; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

OY 2 LVWGEFQVDRSGARGCQTSNDAAAEALNASSKEALQIREKIPERYLFQFHNLSALA 61
111 LLEGTQIDGSRARGCQQTANDAYAKATNAGSKSQALVDIKELAPROYLVLFHNINSILD 170

OY 62 AFDKTPPE 70
DB 171 KVFQVPPAP 179

RESULT 5
US-07-745-206A-2
; Sequence 2, Application US/07745206A
; Patent No. 5792846
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feltman, Daniel
; APPLICANT: McCue, Ann
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: Human Calcium Channel Compositions and
; Methods of Screening
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fitch, Even, Tabin & Flannery
; STREET: 135 S. LaSalle
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:

```

```

1  APPLICATION NUMBER: US/07/745,206A
2  FILING DATE: 19910815
3  CLASSIFICATION: 435
4  ATTORNEY/AGENT INFORMATION:
5  NAME: ROBER, Scott B.
6  REFERENCE/DOCKET NUMBER: 51504
7  TELECOMMUNICATION INFORMATION:
8  TELEPHONE: 312-372-7842
9  INFORMATION FOR SEQ ID NO: 2:
10 SEQUENCE CHARACTERISTICS:
11 LENGTH: 2161 amino acids
12 TYPE: AMINO ACID
13 TOPOLOGY: linear
14 MOLECULE TYPE: protein
15 US-07-745-206A-2

Query Match 17.7%; Score 63.5; DB 1; Length 2161;
Best Local Similarity 28.4%; Pred. No. 15;
Matches 21; Conservative 5; Mismatches 19; Indels 29; Gaps 1;

OY 4 WGEFQVDRSGARGCQTS-----NDAAEALNASSK 34
DB 707 WNAVVDGIMAYGCPSSCMVICYIFILFCGNYLLWFLAVLADNLAESLNTAQK 766

OY 35 EBAIQIIRKIPKEK 48
111 : : : : :
DB 767 EAEKEKRIKARK 780

RESULT 6
US-08-455-543A-49
; Sequence 49, Application US/08455543A
; Patent No. 5792846
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feltman, Daniel
; APPLICANT: McCue, Ann
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; METHODS OF SCREENING
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,543A
; FILING DATE: May 31, 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/223,305
; FILING DATE: April 4, 1994
; PRIOR APPLICATION NUMBER: 07/868,354
; FILING DATE: April 10, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/745,206
; FILING DATE: 15-SEP-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/620,250
; FILING DATE: 30-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/482,384

```

```

? FILING DATE: 20-FEB-1990
? APPLICATION DATA:
? APPLICATION NUMBER: US 07/603,751
? PRIOR APPLICATION DATA:
? PRIOR APPLICATION NUMBER: 07/176,899
? FILING DATE: 04-APR-1989
? APPLICATION DATA:
? APPLICATION NUMBER: WO PCT/US89/01408
? FILING DATE: 04-APR-1989
? APPLICATION DATA:
? APPLICATION NUMBER: US 07/176,899
? FILING DATE: 04-APR-1989
? ATTORNEY/AGENT INFORMATION:
? NAME: Seligman, Stephanie L.
? REGISTRATION NUMBER: 33,779
? REFERENCE/DOCKET NUMBER: 6362-52517
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (619)238-0999
? TELEFAX: (619)238-0062
? INFORMATION FOR SEQ ID NO: 49:
? SEQUENCE CHARACTERISTICS:
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? FRAGMENT TYPE: Internal
?
US-08-455-543A-49

Query Watch 17.7% Score 63.5; DB 1; Length 2161;
Best Local Similarity 28.4%; Pred. No. 15;
Matches 21; Conservative 5; Mismatches 19; Indels 29; Gaps 1;

QY 4 MGEFQVDRSGARGCQTS-----NDAAAEALNASK 34
DB 707 NNAYWDCIMAYGSSSCHMVCIVIFILICGNVILLNVLATVDNLADRESLNTAOK 766

QY 35 EEAQLIIEKPEK 48
DB 767 BEAEKERKKIARK 780

US-08-455-543A-51

Query Watch 17.7% Score 63.5; DB 1; Length 2161;
Best Local Similarity 28.4%; Pred. No. 15;
Matches 21; Conservative 5; Mismatches 19; Indels 29; Gaps 1;

QY 4 MGEFQVDRSGARGCQTS-----NDAAAEALNASK 34
DB 707 NNAYWDCIMAYGSSSCHMVCIVIFILICGNVILLNVLATVDNLADRESLNTAOK 766

QY 35 EEAQLIIEKPEK 48
DB 767 BEAEKERKKIARK 780

US-08-455-543A-51

RESULT 7
US-08-455-543A-51
: Sequence 51, Application US/08455543A
: Patent No. 5792845
: GENERAL INFORMATION:
: APPLICANT: Harpold, Michael
: APPLICANT: Ellis, Steven
: APPLICANT: Williams, Mark
: APPLICANT: Feldman, Daniel
: APPLICANT: Brenner, Robert
: TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
: TITLE OF SEQUENCES: METHODS
: NUMBER OF SEQUENCES: 57
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Brown, Martin, Haller & McClain
: STREET: 1660 Union Street
: CITY: San Diego
: STATE: California
: COUNTRY: USA
: ZIP: 92101-2926
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: SOFTWARE: FASTSEQ Version 1.5
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/455,543A
: FILING DATE: May 31, 1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/223-305
: FILING DATE: April 4, 1994
```

```

? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 07/868,354
? FILING DATE: April 10, 1992
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/745,206
? FILING DATE: 15-AUG-1991
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/620,250
? FILING DATE: 30-NOV-1990
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/482,384
? FILING DATE: 20-FEB-1990
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/603,751
? FILING DATE: 04-APR-1989
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: WO PCT/US89/01408
? FILING DATE: 04-APR-1989
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/176,899
? FILING DATE: 04-APR-1989
? ATTORNEY/AGENT INFORMATION:
? NAME: Seligman, Stephanie L.
? REGISTRATION NUMBER: 33,779
? REFERENCE/DOCKET NUMBER: 6362-52517
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (619)238-0999
? TELEFAX: (619)238-0062
? INFORMATION FOR SEQ ID NO: 51:
? SEQUENCE CHARACTERISTICS:
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? FRAGMENT TYPE: Internal
?
US-08-455-543A-51

Query Watch 17.7% Score 63.5; DB 1; Length 2161;
Best Local Similarity 28.4%; Pred. No. 15;
Matches 21; Conservative 5; Mismatches 19; Indels 29; Gaps 1;

QY 4 MGEFQVDRSGARGCQTS-----NDAAAEALNASK 34
DB 707 NNAYWDCIMAYGSSSCHMVCIVIFILICGNVILLNVLATVDNLADRESLNTAOK 766

QY 35 EEAQLIIEKPEK 48
DB 767 BEAEKERKKIARK 780

US-08-455-543A-51

RESULT 8
US-08-223-305C-49
: Sequence 49, Application US/08223305C
: Patent No. 5851824
: GENERAL INFORMATION:
: APPLICANT: Harpold, Michael
: APPLICANT: Ellis, Steven
: APPLICANT: Williams, Mark
: APPLICANT: Feldman, Daniel
: APPLICANT: McCue, Ann
: APPLICANT: Brenner, Robert
: TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
: TITLE OF SEQUENCES: METHODS
: NUMBER OF SEQUENCES: 57
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Brown, Martin, Haller & McClain
: STREET: 1660 Union Street
: CITY: San Diego
: STATE: California
: COUNTRY: USA
: ZIP: 92101-2926
```

```

1  COMPUTER READABLE FORM:
2  MEDIUM TYPE: Diskette
3  COMPUTER: IBM Compatible
4  OPERATING SYSTEM: DOS
5  SOURCE: FSI SEQ Ion 1.5
6  CURRENT APPLICATION DATA:
7  APPLICATION NUMBER: US/08/223.305C
8  FILING DATE: April 4, 1994
9  PRIOR APPLICATION DATA:
10 APPLICATION NUMBER: 07/868.354
11 FILING DATE: April 10, 1992
12 PRIOR APPLICATION DATA:
13 APPLICATION NUMBER: US/07/745.206
14 FILING DATE: 15-AUG-1991
15 PRIOR APPLICATION DATA:
16 APPLICATION NUMBER: US/07/620.250
17 FILING DATE: 30-NOV-1990
18 PRIOR APPLICATION DATA:
19 APPLICATION NUMBER: US/07/482.384
20 FILING DATE: 20-FEB-1990
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER: US/07/603.751
23 FILING DATE: 04-APR-1989
24 NAME: Seidman, Stephanie L.
25 REGISTRATION NUMBER: 33,779
26 REFERENCE/DOCKET NUMBER: 52516 (P519739)
27 TELECOMMUNICATION INFORMATION:
28 TELEPHONE: (619)238-0999
29 TELEFAX: (619)238-0624
30 INFORMATION FOR SEQ ID NO: 49:
31 SEQUENCE CHARACTERISTICS:
32 LENGTH: 2161 amino acids
33 TYPE: amino acid
34 STRANDEDNESS: single
35 TOPOLOGY: linear
36 MOLECULE TYPE: protein
37 FRAGMENT TYPE: Internal
38 US-08-223-305C-49
39
40 Query Match 17.7%; Score 63.5; DB 2; Length 2161:
41 Best Local Similarity 28.4%; Pred. No. 15;
42 Matches 21; Conservative 5; Mismatches 19; Indels 29; Gaps 1:
43
44 QY 4 WGFQVQDSRGRCQTS-----NDAAEALNASK 34
45 DB 707 WNAVWDGIMAYGPPSSGNCVIFILFGNVTLLNVFLAIVADLADSLNTAQK 766
46
47 QY 35 EEAQIREKIPEK 48
48 DB 767 EEAEKERKTIARK 780
49
50 RESULT 9
51 US-08-223-305C-51
52 Sequence 51, Application US/08223305C
53 Patent No. 5851824
54 GENERAL INFORMATION:
55 APPLICANT: Harpold, Michael
56 APPLICANT: Williams, Mark
57 APPLICANT: Feldman, Daniel
58 APPLICANT: McCue, Ann
59 APPLICANT: Brenner, Robert
60 TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
61 METHODS

```

```

1  NUMBER OF SEQUENCES: 57
2  CORRESPONDENCE ADDRESS:
3  ADDRESSEE: Brown, Martin, Haller & McClain
4  STREET: 1660 Union Street
5  CITY: San Diego
6  STATE: California
7  COUNTRY: USA
8  ZIP: 92101-2926
9  COMPUTER READABLE FORM:
10 MEDIUM TYPE: Diskette
11 COMPUTER: IBM Compatible
12 OPERATING SYSTEM: DOS
13 SOURCE: FSI SEQ Ion 1.5
14 CURRENT APPLICATION DATA:
15 APPLICATION NUMBER: US/08/223.305C
16 FILING DATE: April 4, 1994
17 PRIOR APPLICATION DATA:
18 APPLICATION NUMBER: 07/868.354
19 FILING DATE: April 10, 1992
20 PRIOR APPLICATION DATA:
21 APPLICATION NUMBER: US/07/745.206
22 FILING DATE: 15-AUG-1991
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: US/07/620.250
25 FILING DATE: 30-NOV-1990
26 PRIOR APPLICATION DATA:
27 APPLICATION NUMBER: US/07/482.384
28 FILING DATE: 20-FEB-1990
29 PRIOR APPLICATION DATA:
30 APPLICATION NUMBER: US/07/603.751
31 FILING DATE: 04-APR-1989
32 NAME: Seidman, Stephanie L.
33 REGISTRATION NUMBER: 33,779
34 REFERENCE/DOCKET NUMBER: 52516 (P519739)
35 TELECOMMUNICATION INFORMATION:
36 TELEPHONE: (619)238-0999
37 TELEFAX: (619)238-0624
38 INFORMATION FOR SEQ ID NO: 51:
39 SEQUENCE CHARACTERISTICS:
40 LENGTH: 2161 amino acids
41 TYPE: amino acid
42 STRANDEDNESS: single
43 TOPOLOGY: linear
44 MOLECULE TYPE: protein
45 FRAGMENT TYPE: Internal
46 US-08-223-305C-51
47
48 Query Match 17.7%; Score 63.5; DB 2; Length 2161:
49 Best Local Similarity 28.4%; Pred. No. 15;
50 Matches 21; Conservative 5; Mismatches 19; Indels 29; Gaps 1:
51
52 QY 4 WGFQVQDSRGRCQTS-----NDAAEALNASK 34
53 DB 707 WNAVWDGIMAYGPPSSGNCVIFILFGNVTLLNVFLAIVADLADSLNTAQK 766
54
55 QY 35 EEAQIREKIPEK 48
56 DB 767 EEAEKERKTIARK 780
57
58 RESULT 10
59 US-08-311-363-2
60 Sequence 2, Application US/08311363
61 Patent No. 5876958
62 GENERAL INFORMATION:

```

```

? APPLICANT: Harpold, Michael
? APPLICANT: Ellis, Steven
? APPLICANT: Williams, Mark
? APPLICANT: McLaman, Daniel
? APPLICANT: McLaman, Daniel
? APPLICANT: Brenner, Robert
? TITLE OF INVENTION: Human Calcium Channel Compositions and
? TITLE OF INVENTION: Methods
? NUMBER OF SEQUENCES: 32
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Brown, Martin, Haller & McClain
? STREET: 1660 Union Street
? CITY: San Francisco
? STATE: California
? COUNTRY: USA
? ZIP: 92101-2926
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? SOFTWARE: PC-DOS/MS-DOS
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/311,363
? FILING DATE:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/745,206
? FILING DATE: 15 AUG 1991
? ATTORNEY/AGENT INFORMATION:
? NAME: Seidman, Stephanie L.
? REGISTRATION NUMBER: 33,779
? REFERENCE/DOCKET NUMBER: 6362-51506
? TELEPHONE: (619)238-0999
? TELEFAX: (619)238-0062
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2161 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-311-363-2

```

```

Query Match 17.7% Score 63.5; DB 2; Length 2161;
Best Local Similarity 28.4% Pred. No. 15;
Matches 21: Conservative 5; Mismatches 19; Indels 29; Gaps 1;

```

```

QY 4 MGEFQVDRGARGCCOTS-----NDAAAEALNASKK 34
Db 707 MNWVFQDMATWGGPSSGMVCIYFLLFCGNYLLWFLAVDNLDAESLNTAK 766
QY 35 EEAQITREKPEK 48
Db 767 EEAEEKENKKIARK 780

```

```

RESULT 11
US-08-311-692-51
? Sequence 51: Application US/08926692
? Patent No. 5958727
? GENERAL INFORMATION:
? APPLICANT: Brody, Howard
? APPLICANT: Yaver, Deborah S.
? APPLICANT: Lamsa, Michael
? APPLICANT: Hansen, Kim
? TITLE OF INVENTION: Methods for Modifying the Production of
? NUMBER OF SEQUENCES: 80
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: No. 5958727o No. 5958727th America, Inc.
? STREET: 405 Lexington Avenue
? CITY: New York
? STATE: NY

```

```

? COUNTRY: USA
? ZIP: 10174
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? SOFTWARE: PC-DOS/MS-DOS
? OPERATING SYSTEM: IBM compatible
? SOFTWARE: FASTISO for Windows Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/928,692
? FILING DATE: 12-SEPT-1997
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Hendricks, Glenna
? REGISTRATION NUMBER: 32,595
? REFERENCE/DOCKET NUMBER: 4944-200-US
? TELEPHONE: 212-867-0123
? TELEFAX: 212-878-9655
? INFORMATION FOR SEQ ID NO: 51:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2161 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: No. 5958727e
? US-08-328-692-51

```

```

Query Match 16.2% Score 59; DB 2; Length 374;
Best Local Similarity 35.3% Pred. No. 8.3;
Matches 18: Conservative 5; Mismatches 12; Indels 2;

```

```

QY 22 NDAAAEALNASKK--EEAQITREKPEKYLGFQHNINSALAAIFDKTPE 69
Db 181 SDAGRELAUTASAKAGLEVLQLIHPV-----AAALAYDARPE 218

```

```

RESULT 12
US-08-446-100-26
? Sequence 26: Application US/08446100
? Patent No. 6001625
? GENERAL INFORMATION:
? APPLICANT: Bloomfield, Clarence A
? APPLICANT: Bloomfield, Michael B
? APPLICANT: Lockridge, Oksana
? TITLE OF INVENTION: Site-Directed Mutagenesis of Esterases
? NUMBER OF SEQUENCES: 31
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Hendricks and Assoc.
? STREET: 9669 A Main Street, P.O. Box 2509
? CITY: Fairfax
? STATE: Virginia
? COUNTRY: US
? ZIP: 22031
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: FASTISO for Windows Version 1.0, Version 11.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/446,100
? FILING DATE: 19-MAY-1995
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Hendricks, Glenna
? REGISTRATION NUMBER: 32,595
? REFERENCE/DOCKET NUMBER: 4944-200-US
? TELEPHONE: (703) 425-4250
? TELEFAX: (703) 425-2767
? INFORMATION FOR SEQ ID NO: 26:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 454 amino acids

```



```

: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: HYPOTHETICAL: YES
: ANTI-SENSE: YES
: FRAGMENT TYPE: N-terminal
: ORIGINAL SOURCE:
: ORGANISM: human esterases
:
US-08-446-100-26

Query Match 15.9%: Score 57: DB 3: Length 454;
Best Local Similarity 30.6%: Pred. No. 15;
Matches 19: Conservative 10; Mismatches 23; Indels 10; Gaps 4:

OY 17 GCOTSDAA-AEALNASKEALQIREKIPKYLQFPHNL---NSALAAIFD-----KT 67
DB 160 GKTTTSAANYHCLQKTEELLETTL-KIGNSYLWTRQRETLTGTVIDGMILLKLT 218
OY 68 PE 69
DB 219 PE 220

RESULT 14
US-08-446-100-28
: Sequence 28, Application US/08446100
: Patent No. 6001625
: GENERAL INFORMATION:
: APPLICANT: Broomfield, Clarence A
: APPLICANT: Millard, Charles B
: APPLICANT: Lockridge, Oksana
: TITLE OF INVENTION: Site-Directed Mutagenesis of Esterases
: NUMBER OF SEQUENCES: 31
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Hendricks and Assoc.
: STREET: 9669 A Main Street, P.O. Box 2509
: CITY: Fairfax
: STATE: VA
: COUNTRY: US
: ZIP: 22031
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/446,100
: FILING DATE: 19-MAY-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Hendricks, Glenna
: REGISTRATION NUMBER: 32,535
: REFERENCE/DOCKET NUMBER: broomfield
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 425-4250
: TELEFAX: (703) 425-2767
: INFORMATION FOR SEQ ID NO: 28:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 454 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: HYPOTHETICAL: YES
: ANTI-SENSE: YES
: FRAGMENT TYPE: N-terminal
: ORIGINAL SOURCE:
: ORGANISM: human esterases
:
US-08-446-100-28

Query Match 15.9%: Score 57: DB 3: Length 454;
Best Local Similarity 30.6%: Pred. No. 15;
Matches 19: Conservative 10; Mismatches 23; Indels 10; Gaps 4:

OY 17 GCOTSDAA-AEALNASKEALQIREKIPKYLQFPHNL---NSALAAIFD-----KT 67
DB 160 GKTTTSAANYHCLQKTEELLETTL-KIGNSYLWTRQRETLTGTVIDGMILLKLT 218
OY 68 PE 69
DB 219 PE 220

US-08-446-100-27

Query Match 15.9%: Score 57: DB 3: Length 454;
Best Local Similarity 30.6%: Pred. No. 15;
Matches 19: Conservative 10; Mismatches 23; Indels 10; Gaps 4:

OY 17 GCOTSDAA-AEALNASKEALQIREKIPKYLQFPHNL---NSALAAIFD-----KT 67
DB 160 GKTTTSAANYHCLQKTEELLETTL-KIGNSYLWTRQRETLTGTVIDGMILLKLT 218
OY 68 PE 69
DB 219 PE 220

RESULT 13
US-08-446-100-27
: Sequence 27, Application US/08446100
: Patent No. 6001625
: GENERAL INFORMATION:
: APPLICANT: Broomfield, Clarence A
: APPLICANT: Millard, Charles B
: APPLICANT: Lockridge, Oksana
: TITLE OF INVENTION: Site-Directed Mutagenesis of Esterases
: NUMBER OF SEQUENCES: 31
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Hendricks and Assoc.
: STREET: 9669 A Main Street, P.O. Box 2509
: CITY: Fairfax
: STATE: VA
: COUNTRY: US
: ZIP: 22031
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/446,100
: FILING DATE: 19-MAY-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Hendricks, Glenna
: REGISTRATION NUMBER: 32,535
: REFERENCE/DOCKET NUMBER: broomfield
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 425-4250
: TELEFAX: (703) 425-2767
: INFORMATION FOR SEQ ID NO: 27:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 454 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: HYPOTHETICAL: YES
: ANTI-SENSE: YES
: FRAGMENT TYPE: N-terminal
: ORIGINAL SOURCE:
: ORGANISM: human esterases
:
US-08-446-100-27
```

```

RESULT 15
US-08-446-100-29 Application US/08446100
; Sequence 29, 60022
; PRTS: 1, 60022
; GENERAL INFORMATION:
; APPLICANT: Broomfield, Clarence A
; APPLICANT: Millard, Charles B
; APPLICANT: Lockridge, Oksana
; TITLE OF INVENTION: Site-Directed Mutagenesis of Esterases
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Broomfield and Assoc.
; STREET: 9669 A Main Street, P.O. Box 2509
; CITY: Fairfax
; STATE: VA
; COUNTRY: US
; ZIP: 22031
; COMPUTER READABLE FORM:
; 1. Copy to disk
; 2. Submit to file
; 3. Computer: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,100
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hendricks Glenna
; REGISTRATION NUMBER: 32,535
; REFERENCE/DOCKET NUMBER: Broomfield
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 425-4250
; TELEFAX: (703) 425-2767
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 454 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; IMMEDIATE SOURCE: YES
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: human esterases
;
US-08-446-100-29

```

```

Query Match 15.94; Score 57; DB 3; Length 454;
Local Similarity 10.64; Pref. No 13;
Matches 19; Conservative 10; Mismatches 23; Indels 10; Gaps 4;

Qy 17 GCOTSDAA-ASALNASSKEALQIREKIPEYLFQFHNL---NSALAAFD----KT 67
Db 160 CKTTTSAANVCLCRKTEELLETTL-KIGNSYLWYHRETFQESTLLGTVDICMLLKT 218
Qy 68 PF 69
Db 219 PF 220

```

Search completed: February 3, 2001, 02:19:48  
Job time: 1710 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:22:52 ; Search time 180.33 Seconds  
(without alignments)  
45.497 Million cell updates/sec

Title: US-09-289-346a-8

Perfect score: 351

Sequence: 1 TLWGEAAVDCRSRGCCOT.....FQPHNLSNLDRIFFKTPPE 70

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL15:\*
- 1: sp.archaea.\*
- 2: sp.bacteria.\*
- 3: sp.fungi.\*
- 4: sp.human.\*
- 5: sp.invertebrate.\*
- 6: sp.mammal.\*
- 7: sp.mhc.\*
- 8: sp.organelle.\*
- 9: sp.phage.\*
- 10: sp.plant.\*
- 11: sp.rodent.\*
- 12: sp.virus.\*
- 13: sp.vertabrate.\*
- 14: sp.unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	DB ID	Description
1	282	78.1	226	12	009727	009727 leonurus mo
2	282	78.1	226	12	09WH6	09WH6 tomato mild
3	280	77.6	361	12	Q67574	Q67574 bean golden
4	278	77.0	225	12	Q9Q081	Q9Q081 cowpea gold
5	275	76.9	185	12	Q9Q081	Q9Q081 sweet potato
6	275	76.9	185	12	Q9Q081	Q9Q081 sweet potato
7	267	74.0	149	12	P88975	P88975 macroptiliu
8	267	74.0	233	12	Q9YLA4	Q9YLA4 macroptiliu
9	261	72.3	234	12	Q93180	Q93180 potato yell
10	258	71.5	190	12	Q92089	Q92089 tobacco lea
11	258	71.5	190	12	Q92084	Q92084 tobacco lea
12	257	71.2	190	12	Q92087	Q92087 tobacco lea
13	255	70.6	208	12	Q92084	Q92084 tobacco lea
14	255	70.6	208	12	Q92087	Q92087 tobacco lea
15	255	70.6	208	12	Q92086	Q92086 tobacco lea
16	253	70.1	208	12	Q92086	Q92086 tobacco lea
17	252	69.8	203	12	Q92083	Q92083 tobacco lea
18	252	69.8	363	12	Q73577	Q73577 cotton leaf
19	245	67.9	208	12	Q920A0	Q920A0 tobacco lea

20	245	67.9	363	12	072705	072705 cotton leaf
21	245	67.9	363	12	072719	072719 cotton leaf
22	244	67.6	190	12	092086	092086 tobacco lea
23	244	67.6	190	12	092086	092086 tobacco lea
24	243	67.3	362	12	056816	056816 chayote pec
25	243	67.3	362	12	056816	056816 chayote pec
26	242	67.0	359	12	09Y2V4	09Y2V4 tomato yell
27	242	67.0	359	12	09Y2V2	09Y2V2 tomato yell
28	242	67.0	359	12	09YUX7	09YUX7 tomato yell
29	242	67.0	359	12	09YUX7	09YUX7 tomato yell
30	241	66.8	353	12	072692	072692 beet curly
31	241	66.8	353	12	072692	072692 beet curly
32	241	66.8	353	12	088435	088435 tobacco lea
33	240	66.5	100	12	092NA7	092NA7 tobacco lea
34	240	66.5	362	12	091NA8	091NA8 okra enatio
35	239	66.2	307	12	091ET7	091ET7 cotton leaf
36	239	66.2	361	12	072723	072723 cotton leaf
37	238	65.9	231	12	096620	096620 african tom
38	238	65.9	354	12	091NA2	091NA2 south afric
39	237	65.7	358	12	092NA5	092NA5 okra enatio
40	237	65.7	358	12	092NA5	092NA5 okra enatio
41	233	64.5	362	12	09YVT7	09YVT7 alchea rose
42	233	64.5	363	12	073494	073494 okra yellow
43	232	64.3	351	12	09Q9R3	09Q9R3 dieliptera
44	231	64.0	306	12	091ET1	091ET1 cotton leaf
45	231	64.0	358	12	09WR17	09WR17 african cas

ALIGNMENTS

```
RESULT 1
009727 PRELIMINARY; PRT; 226 AA.
AC 009727: 1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last annotation update)
ID 009727: 1997 (TrEMBLrel. 04, Last annotation update)
DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE REP PROTEIN (FRAGMENT).
GN REP.
OS Leonurus mesale virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=58177;
RA STRAIN=LEWV- BRAZIL 1;
RC STRAIN=LEWV- BRAZIL 1;
RL Submitted (MAR-1997) to the EMBL/Genbank/DDBJ databases.
DR EMBL; U92532; AAB51157.1;
DR INTERPHO; IPRO01191; -.
DR INTERPRO; IPR001301;
DR PROSITE; PS00707; GEMCOTALL;
DR PRINTS; PR00227; GEMCOTALL;
DR PRINTS; PR00228; GEMCOTALL;
FT NON_TER 226
SQ SEQUENCE 226 AA; 25617 MW; 73C0B6E76083FC5 CRC64;
Query Match: 78.1%; Score 282; DB 12; Length 226;
Sequence Similarity: 78.6%; Pct Ident: 54.2;
Matches 55; Conservative 4; Mismatches 11; Indels 0; Gaps 0;
```

```
QY 1 TLWGEAAVDCRSRGCCOTSNDAAEALQITREKTPKYLFIQPHNLSNL 60
: ||| |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 111 TVENGEEQVDCRSRGCCQTVNDAAAEALNAPDKRTALQITREKTPKYLFIQPHNLSNL 170
QY 61 DRIFKTPPE 70
DB 171 DRIFKAKPEP 180
```

RESULT 2  
Q9WH6

```

ID Q9WHF6 PRELIMINARY: PRT: 226 AA.
AC Q9WHF6;
DT 01-NOV-1999 (TEMBLrel. 12, Created)
DT 01-NOV-1999 (TEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TEMBLrel. 14, Last annotation update)
DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN REP.
OS tomato mild mottle geminivirus.
OC Viruses: ssDNA viruses: Geminiviridae; Unclassified Geminiviridae.
OX NCBI_TaxID=92943;
RN SEQUENCE FROM N.A.
RA Faria J.C.;
RA Nakha W.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,
RA Nakha W.K.
RT "Molecular characterization and DNA-based detection methods for
RT vegetable-infecting geminiviruses in Central America.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF131071; AA033471.1; -
DR INTERPRO: IPRO01191; -
DR PRINTS: PR00227; GEMCOATL1.
DR PFAM: PF00799; Gemin_A1; 1.
DR PRINTS: PR00227; GEMCOATL1.
DR NON_TER 226 226
SQ SEQUENCE 226 AA: 25941 MW: 2EA4116712871A23 CRC64;

Query Match 79.1%; Score 292; DB 12; Length 226;
Best Local Similarity 74.3%; Pred. No. 5, 4e-23;
Matches 52; Conservative 12; Mismatches 6; Indels 0; Gaps 0;

QY 1 TLWGEAANDGSRGCGCOTSDNDAAEALNASSKEEAQIIRKIPKYLFOFHNLSNL 60
Db 111 TLMGEFOTDGRSAGGQOOTANDAAEALNASSKEEAQIIRKIPKYLFOFHNLSNL 170
QY 61 DRIFKTPPEP 70
Db 171 DRIFAKAPEP 180

RESULT 3
T67574 PRELIMINARY: PRT: 361 AA.
AC Q67574;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-JUN-2000 (TEMBLrel. 14, Last annotation update)
DE PUTATIVE REPLICATIVE PROTEIN.
GN AL1.
OS bean golden mosaic virus.
OC Viruses: ssDNA viruses: Geminiviridae; Begomovirus.
OX NCBI_TaxID=10839;
RN SEQUENCE FROM N.A.
RA Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;
RA Maxwell D.P., Russell D.R.;
RL Phytopathology 81:960-965(1991).
RN SEQUENCE FROM N.A.
RA Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,
RA Morales F.J., Maxwell D.P.;
RL Plant Dis. 75:336-342(1991).
RN SEQUENCE FROM N.A.
RA Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;
RA Maxwell D.P., Russell D.R.;
RL EMBL: M89656; AA446312.1; -
DR INTERPRO: IPRO01191; -
DR PRINTS: PR001301; -
DR PFAM: PF00799; Gemin_A1; 1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00227; GEMCOATL1.

```

```

DR PRODM: PD000736; -; 1.
SQ SEQUENCE 361 AA: 41041 MW: 0094C7ACAFD6B78B CRC64;

Query Match 77.6%; Score 280; DB 12; Length 361;
Best Local Similarity 77.6%; Pred. No. 1, 5e-22;
Matches 52; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 4 MGEAANDGSRGCGCOTSDNDAAEALNASSKEEAQIIRKIPKYLFOFHNLSNLDRI 63
Db 113 NGHFQVODGRSAGGQOOTANDAAEALNASSKEEAQIIRKIPKYLFOFHNLSNLDRI 172
QY 64 FKXTPPEP 70
Db 173 FKXAPDP 179

RESULT 4
Q9QDB1 PRELIMINARY: PRT: 225 AA.
ID Q9QDB1;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TEMBLrel. 14, Last annotation update)
DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN REP.
OS cowpea golden mosaic geminivirus.
OC Viruses: ssDNA viruses: Geminiviridae; Begomovirus.
OX NCBI_TaxID=62203;
RN SEQUENCE FROM N.A.
RA Faria J.C.;
RA "Partial nucleotide sequence of cowpea golden mosaic geminivirus from
RT Brazil.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF131071; AA033471.1; -
DR INTERPRO: IPRO01191; -
DR PFAM: PF00799; Gemin_A1; 1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00228; GEMCOATL1.
DR NON_TER 225 225
SQ SEQUENCE 225 AA: 25766 MW: 1089CB6B6D15B5D CRC64;

Query Match 77.0%; Score 278; DB 12; Length 225;
Best Local Similarity 77.6%; Pred. No. 1, 5e-22;
Matches 52; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 4 MGEAANDGSRGCGCOTSDNDAAEALNASSKEEAQIIRKIPKYLFOFHNLSNLDRI 63
Db 113 NGHFQVODGRSAGGQOOTANDAAEALNASSKEEAQIIRKIPKYLFOFHNLSNLDRI 172
QY 64 FKXTPPEP 70
Db 173 FKXPPPEP 179

RESULT 5
Q9Q855 PRELIMINARY: PRT: 364 AA.
ID Q9Q855;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TEMBLrel. 15, Last annotation update)
DE APPLICATION INITIATION PROTEIN AC1.
GN REP.
OS sweet potato leaf curl virus.
OC Viruses: ssDNA viruses: Geminiviridae; Begomovirus.
OX NCBI_TaxID=100755;
RN SEQUENCE FROM N.A.

```

```

RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RP "Detection of a geminivirus infecting sweet potato in the United
RT States"; 82:1253-1257(1998).
RL NCBI_TaxID=104036; AAD4717.1; -.
DR EMBL: AF04036; AAD4717.1; -.
DR INTERPRO: IPR001191; -.
DR INTERPRO: IPR001191; -.
DR PFAM: PF00799; Gemini_AL1; 1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00228; GEMCOATCLVL1.
SQ SEQUENCE 364 AA; 40680 MW; 5F79752431A09D6E CRC64;

Query Match 77.0%; Score 278; DB 12; Length 364;
Best Local Similarity 80.9%; Pred. No. 2.5e-22;
Matches 55; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy 1 TLVNGEAVDGRARGCGQTSNDAAAEALNASKEEALQIIRKIPKYLFOFHNLSNL 60
Db 110 TIENGVFQDGRSARGCGQTSNDAAAEALNASKEEALQIIRKIPKYLFOFHNLSNL 169

Qy 61 DRIFDKTP 68
Db 170 DRIFSKPP 177

RESULT 6
Q98593 PRELIMINARY; PRT; 185 AA.
AC Q98603;
DT 01-FEB-1997 (TrEMBLrel. 02, Created).
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE REP PROTEIN (FRAGMENT).
OS Macrotellium golden mosaic geminivirus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=51034;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAMAICA;
RA Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;
RL Molecular characterization of two distinct geminiviruses infecting M.
EMBL: U67926; AAB97462.1; -.
DR INTERPRO: IPR001191; -.
DR INTERPRO: IPR001191; -.
DR PFAM: PF00799; Gemini_AL1; 1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00228; GEMCOATCLVL1.
FT NON_TER 185 185
FT NON_TER 185 185
SQ SEQUENCE 185 AA; 20975 MW; 3913850A025A5EE1 CRC64;

Query Match 76.2%; Score 275; DB 12; Length 185;
Best Local Similarity 71.4%; Pred. No. 2.5e-22;
Matches 50; Conservative 12; Mismatches 8; Indels 0; Gaps 0;

Qy 1 TLVNGEAVDGRARGCGQTSNDAAAEALNASKEEALQIIRKIPKYLFOFHNLSNL 60
Db 89 TIENGVFQDGRSARGCGQTSNDAAAEALNASKEEALQIIRKIPKYLFOFHNLSNL 148

Qy 61 DRIFDKTP 70
Db 149 DRIFSKPP 158

RESULT 7
P88975 PRELIMINARY; PRT; 149 AA.
AC P88975;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)

```

```

DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
DR Macrotellium golden mosaic geminivirus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=51676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAMAICA;
RA Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;
RL Molecular characterization of two distinct geminiviruses infecting M.
EMBL: U67926; AAB97462.1; -.
DR INTERPRO: IPR001191; -.
DR PFAM: PF00799; Gemini_AL1; 1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00228; GEMCOATCLVL1.
FT NON_TER 149 149
FT NON_TER 149 149
SQ SEQUENCE 149 AA; 16785 MW; E4CF5ED4C9CD508 CRC64;

Query Match 74.0%; Score 267; DB 12; Length 149;
Best Local Similarity 70.0%; Pred. No. 1.4e-21;
Matches 49; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

Qy 1 TLVNGEAVDGRARGCGQTSNDAAAEALNASKEEALQIIRKIPKYLFOFHNLSNL 60
Db 52 TIENGVFQDGRSARGCGQTSNDAAAEALNASKEEALQIIRKIPKYLFOFHNLSNL 111

Qy 61 DRIFDKTP 70
Db 112 DRIFKDPPEP 121

RESULT 8
Q9YLAA PRELIMINARY; PRT; 233 AA.
AC Q9YLAA;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
OS Macrotellium golden mosaic geminivirus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=51676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAMAICA STRAIN 1;
RA Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;
RL Molecular characterization of two distinct geminiviruses infecting M.
EMBL: U67926; AAB97462.1; -.
DR INTERPRO: IPR001191; -.
DR PFAM: PF00799; Gemini_AL1; 1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00228; GEMCOATCLVL1.
FT NON_TER 233 233
SQ SEQUENCE 233 AA; 26355 MW; AA490AF4D2166A02 CRC64;

Query Match 74.0%; Score 267; DB 12; Length 233;
Best Local Similarity 70.0%; Pred. No. 2.4e-21;

```

Matches 49; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

QY 1 TLVNGEAAVDGSRGCGTSDNAAALNASSKEALQITREKIPKYLFOFHNLMSL 60  
 DB 110 TLENGTFOIDGSRGCGTSDNAAALNASSKEALQITREKIPKYLFOFHNLMSL 169  
 QY 61 DRIFDKTPEP 70  
 DB 170 DRIFKMDPEP 179

RESULT 9  
 Q39150 ID Q39180 PRELIMINARY; PRT: 234 AA.  
 AC Q39180;  
 DT 01-JAN-1998 (TEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)  
 DT 01-OCT-2000 (TEMBLrel. 15, Last annotation update)  
 DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).  
 OS Tobacco yellow mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10827;  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-TOMATO STRAIN;  
 RA Guzman P., Arredondo C.R., Emmatty D., Portillo R.J., Gilbertson R.L.;  
 RL Plant Dis. 65:312-314(1997).  
 DR PFAM: PF00799; Gemini\_Ali: 1;  
 DR INTERPRO: IPR001191; -;  
 DR INTERPRO: IPR001191; -;  
 DR PFAM: PF00799; Gemini\_Ali: 1;  
 DR PRINTS: PR00227; GEMCOATCLVL.  
 DR PRINTS: PR00228; GEMCOATCLVL.  
 FT NON-TER 234  
 SO SEQUENCE 234 AA; 26486 MW; 9ED8F0697105CD19 CRC64;

Query Match 72.3%; Score 261; DB 12; Length 234;  
 Best Local Similarity 68.6%; Pred. No. 1.le-20;  
 Matches 48; Conservative 13; Mismatches 9; Indels 0; Gaps 0;

QY 1 TLVNGEAAVDGSRGCGTSDNAAALNASSKEALQITREKIPKYLFOFHNLMSL 60  
 DB 110 TLENGTFOIDGSRGCGTSDNAAALNASSKEALQITREKIPKYLFOFHNLMSL 169  
 QY 61 DRIFDKTPEP 70  
 DB 170 DRIFKMDPEP 179

RESULT 10  
 Q32089 ID Q32089 PRELIMINARY; PRT: 190 AA.  
 AC Q32089;  
 DT 01-MAY-1999 (TEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TEMBLrel. 10, Last sequence update)  
 DT 01-MAY-2000 (TEMBLrel. 13, Last annotation update)  
 DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).  
 OS Tobacco leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=67762;  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-TOMATO STRAIN;  
 RA Guzman P., Arredondo C.R., Emmatty D., Portillo R.J., Gilbertson R.L.;  
 RL Plant Dis. 65:312-314(1997).  
 DR PFAM: PF00799; Gemini\_Ali: 1;  
 DR INTERPRO: IPR001191; -;  
 DR INTERPRO: IPR001191; -;  
 DR PFAM: PF00799; Gemini\_Ali: 1;

DR PRINTS: PR00227; GEMCOATCLVL.  
 DR PRINTS: PR00228; GEMCOATCLVL.  
 FT NON-TER 190  
 SO SEQUENCE 190 AA; 21432 MW; AAC093D1D1610PAD CRC64;

Query Match 71.5%; Score 258; DB 12; Length 190;  
 Best Local Similarity 60.0%; Pred. No. 1.le-20;  
 Matches 51; Conservative 8; Mismatches 10; Indels 16; Gaps 1;

QY 1 TLVNGEAAVDGSRGCGTSDNAAALNASSKEALQITREKIPKYLFOFHNLMSL 60  
 DB 85 TLENGTFOIDGSRGCGTSDNAAALNASSKEALQITREKIPKYLFOFHNLMSL 144  
 QY 61 DRI-----FQKTPR 69  
 DB 145 DRIFAPPLEVFCPTASSFDQVPE 169

RESULT 11  
 Q32084 ID Q32084 PRELIMINARY; PRT: 190 AA.  
 AC Q32084;  
 DT 01-MAY-1999 (TEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TEMBLrel. 10, Last sequence update)  
 DT 01-MAY-2000 (TEMBLrel. 13, Last annotation update)  
 DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).  
 OS Tobacco leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=67762;  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-TOMATO STRAIN;  
 RA Guzman P., Arredondo C.R., Emmatty D., Portillo R.J., Gilbertson R.L.;  
 RL Plant Res. 110:247-257(1997).  
 DR EMBL: AB001318; BAA34039.1; -;  
 DR INTERPRO: IPR001191; -;  
 DR INTERPRO: IPR001191; -;  
 DR PFAM: PF00799; Gemini\_Ali: 1;  
 DR PRINTS: PR00227; GEMCOATCLVL.  
 DR PRINTS: PR00228; GEMCOATCLVL.  
 FT NON-TER 190  
 SO SEQUENCE 190 AA; 21444 MW; AAC1C2943E3F01AD CRC64;

Query Match 71.5%; Score 258; DB 12; Length 190;  
 Best Local Similarity 60.0%; Pred. No. 1.le-20;  
 Matches 51; Conservative 8; Mismatches 10; Indels 16; Gaps 1;

QY 1 TLVNGEAAVDGSRGCGTSDNAAALNASSKEALQITREKIPKYLFOFHNLMSL 60  
 DB 85 TLENGTFOIDGSRGCGTSDNAAALNASSKEALQITREKIPKYLFOFHNLMSL 144  
 QY 61 DRI-----FQKTPR 69  
 DB 145 DRIFAPPLEVFCPTASSFDQVPE 169

RESULT 12  
 Q39827 ID Q39827 PRELIMINARY; PRT: 190 AA.  
 AC Q39827;  
 DT 01-NOV-1999 (TEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TEMBLrel. 12, Last sequence update)  
 DT 01-MAY-2000 (TEMBLrel. 13, Last annotation update)  
 DE C1 PROTEIN (FRAGMENT).  
 OS Tobacco leaf curl virus.

OC Viruses: ssDNA viruses: Geminiviridae; Begomovirus.  
 OX NCBI\_TaxId=67762;  
 RN [1]  
 RC SEQUENCE FROM N.A.

RA Ooi K., Ohshita S., Ishii I., Yahara T.;  
 RT "Molecular phylogeny of geminivirus infecting wild plants in Japan."  
 RL J. Plant Res. 110:247-257(1997).  
 DR EMBL: AB001303; BAA34010.1; -.  
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
 DE INTERPRO: IPR001191; -.  
 DR PFAM: PF00799; Gemin1\_1.  
 DE C1 AND C4 GENES, CLONE ABUR3-1, PARTIAL  
 DE AND COMPLETE CDS (FRAGMENT).  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATL1.  
 FT NON\_TER 1 190 190  
 FT NON\_TER 1 190 190  
 SO SEQUENCE 190 AA; 21444 MW; 93C3742A8EBD7EB CRC64;

Query Match 71.2%; Score 257; DB 12; Length 190;  
 Best Local Similarity 60.0%; Pred. No. 2.5e-20;  
 Matches 51; Conservative 8; Mismatches 10; Indels 16; Gaps 1;  
 QY 1 TLWGCEAVDGRSGCGCTSDNAAAEALNASSKEALQIIRKIPKYLFOYHNLNSL 60  
 Db 85 TLWGCEAVDGRSGCGCTSDNAAAEALNASSKEALQIIRKIPKYLFOYHNLNSL 144  
 QY 61 DRI-----FQKTP 69  
 Db 145 DRIAPFLEVPVGFSSSFQVPE 169

RESULT 13  
 Q92C0C Q92C0C PRELIMINARY; PRT: 208 AA.  
 AC Q92C0C;  
 DT 01-MAY-1999 (TReMBLrel. 10, Created)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
 DE INTERPRO: IPR001191; -.  
 DE C1 AND C4 GENES, CLONE ABUR3-1, PARTIAL  
 DE AND COMPLETE CDS (FRAGMENT).  
 DR PFAM: PF00799; Gemin1\_1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATL1.  
 FT NON\_TER 201 201  
 FT NON\_TER 201 201  
 SO SEQUENCE 208 AA; 23526 MW; 249CC31D8729C72D CRC64;

Query Match 71.2%; Score 257; DB 12; Length 208;  
 Best Local Similarity 60.0%; Pred. No. 2.5e-20;  
 Matches 51; Conservative 8; Mismatches 10; Indels 16; Gaps 1;  
 QY 1 TLWGCEAVDGRSGCGCTSDNAAAEALNASSKEALQIIRKIPKYLFOYHNLNSL 60  
 Db 102 TLWGTFVQDGRSGCGCTSDNAAAEALNASSKEALQIIRKIPKYLFOYHNLNSL 161  
 QY 61 DRI-----FQKTP 69  
 Db 162 DRIAPFLEVPVGFSSSFQVPE 186

RESULT 14  
 Q92C0C Q92C0C PRELIMINARY; PRT: 208 AA.  
 AC Q92C0C;  
 DT 01-MAY-1999 (TReMBLrel. 10, Created)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
 DE INTERPRO: IPR001191; -.  
 DE C1 AND C4 GENES, CLONE ABUR3-1, PARTIAL  
 DE AND COMPLETE CDS (FRAGMENT).  
 DR PFAM: PF00799; Gemin1\_1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATL1.  
 FT NON\_TER 201 201  
 FT NON\_TER 201 201  
 SO SEQUENCE 208 AA; 23486 MW; E301135F799C3DAD CRC64;

Query Match 70.6%; Score 255; DB 12; Length 208;  
 Best Local Similarity 56.8%; Pred. No. 4.1e-20;  
 Matches 50; Conservative 9; Mismatches 10; Indels 16; Gaps 1;  
 QY 1 TLWGCEAVDGRSGCGCTSDNAAAEALNASSKEALQIIRKIPKYLFOYHNLNSL 60  
 Db 102 TLWGTFVQDGRSGCGCTSDNAAAEALNASSKEALQIIRKIPKYLFOYHNLNSL 161  
 QY 61 DRI-----FQKTP 69  
 Db 162 DRIAPFLEVPVGFSSSFQVPE 186  
 RESULT 15  
 Q92C0B Q92C0B PRELIMINARY; PRT: 208 AA.  
 AC Q92C0B;  
 DT 01-MAY-1999 (TReMBLrel. 10, Created)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
 DE INTERPRO: IPR001191; -.  
 DE C1 AND C4 GENES, CLONE ABUR3-1, PARTIAL  
 DE AND COMPLETE CDS (FRAGMENT).  
 DR PFAM: PF00799; Gemin1\_1.  
 DR PRINTS: PR00227; GEMCOATL1.  
 DR PRINTS: PR00228; GEMCOATL1.  
 FT NON\_TER 208 208  
 FT NON\_TER 208 208  
 SO SEQUENCE 208 AA; 23472 MW; 629D00EEF7C9956AA CRC64;

```

Query Match      70.6%; Score 255; DB 12; Length 208;
Seq. Locality 58.8%; Pred. No. 4, 1e-20;
Matches 50; Conservative 9; Mismatches 10; Indels 16; Gaps 1;

Qy 1 TLVNGEAAVDGRSARGGQTSNDAAALNASSKEEALQITREKIPKYLQFHNLNSNL 60
   ||| :||||| :||| :||||| :||| :||||| :||| :||||| :||| :|||||
Db 102 TLENGTFQIDGRSARGGCCNANDACAALNASSKADALATIREKLPKDFIQYHNLNSNL 161

Qy 61 DRI-----FDKYPE 69
   ||| :||||| :||| :||||| :||| :||||| :||| :||||| :||| :|||||
Db 162 DRIFAPPLEVFCPPSSSFQVPE 186

```

Search completed: February 3, 2001, 02:22:52  
Job time: 1889 sec



GenCode version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: February 3, 2001, 02:24:28 ; Search time 83.07 seconds  
(without alignments)  
27.213 Million cell updates/sec

Title: US-09-289-346a-8  
Percent score: 301  
Sequence: 1 TLWGFANVDSRGSGCQT.....LQFHNLNLDLRFDTKTEP 70

Scoring table: BLOSUM62 Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	350	97.0	352	1 VALL_TGMV	P03567 tomato gold
2	250	59.3	361	1 VALL_PMYV	P27258 potato yell
3	237	55.7	358	1 VALL_CUWR	P14992 cassava lat
4	237	55.7	358	1 VALL_CUWR	P14992 cassava lat
5	230	63.7	362	1 VALL_TYLCV	P36750 cassava lat
6	223	61.8	359	1 VALL_TYLCV	P36750 cassava lat
7	219	60.7	349	1 VALL_PMYV	P06923 pepper husa
8	216	59.8	359	1 VALL_TYLCV	P07260 tomato yell
9	214	59.3	353	1 VALL_BGMV	P05175 bean golden
10	213	59.0	355	1 VALL_ABMYV	P21947 abutilon mo
11	212	58.7	357	1 VALL_TYLCV	P27259 tomato yell
12	211	58.6	357	1 VALL_TYLCV	P27259 tomato yell
13	210	58.2	361	1 VALL_TGMV	P06557 beet curly
14	210	58.2	361	1 VALL_TGMV	P06557 beet curly
15	117	32.4	347	1 VALL_SLVC	P29048 squash leaf
16	65.5	17.9	299	1 V175_HRLPU	P16787 homo sapien
17	64.5	17.6	1610	1 CCAD_MESAU	Q92447 mesocricetu
18	63.5	17.6	2161	1 CCAD_HUMAN	Q92447 mesocricetu
19	63.5	17.6	2161	1 CCAD_HUMAN	Q92447 mesocricetu
20	61	16.9	1493	1 CYP1_YEAST	P27732 ratius nov
21	60	16.6	447	1 CUSA_DROME	P12351 tobacco yell
22	59.5	16.5	136	1 Y452_CAREL	P56079 d phosphati
23	59.5	16.5	136	1 Y452_CAREL	P56079 d phosphati
24	59.5	16.5	299	1 V175_HELPY	O62250 caenorhabdi
25	58.5	16.2	297	1 RRP_RABVP	P56112 helicobacte
26	58.5	16.2	334	1 G3P_BACST	P06747 rabies viru
27	58.5	16.2	335	1 G3P_BACCO	P00362 bacillus at
28	58.5	16.2	335	1 G3P_BACCO	P15115 bacillus co
29	58.5	16.2	569	1 YP67_HAEIN	P48926 candida alb
30	58.5	16.2	1070	1 P11B_HUMAN	P42338 homo sapien
31	58	16.1	617	1 YACH_ECOLI	P36682 escherichia
32	58	16.1	874	1 SLAP_ECOLI	P49052 bacillus li
33	57	15.8	200	1 VIP_CHTCK	P48143 gallus gall

34 57 15-8 200 1 VIP\_MELGA P45644 melleagris g  
35 57 15-8 247 1 YCF4\_YEAST P25349 saccharomyc  
36 57 15-8 355 1 CRTB\_RHOSH P54905 rhodobacter  
37 57 15-8 387 1 Y4PE\_RHISH P55015 rhizobium s  
38 57 15-8 403 1 Y4PE\_RHISH P55015 rhizobium s  
39 56.5 15.7 403 1 Y4PE\_RHISH P55015 rhizobium s  
40 56.5 15.7 2190 1 CCAD\_CHICK Q37215 phosphatila  
41 56.5 15.7 2190 1 CCAD\_CHICK Q37215 phosphatila  
42 56 15-5 336 1 Y625\_METJA O73700 gallus gall  
43 56 15-5 513 1 HEMO\_CHICK Q58042 methanococ  
44 56 15-5 536 1 ACFA\_EMENI P18080 gallus gall  
45 55.5 15.4 224 1 SP2R\_BACSU P28298 emericella  
P1586 h c-1-tetra  
P39151 bacillus su

## ALIGNMENTS

RESULT 1  
VALL\_TGMV STANDARD; PRT; 352 AA.  
ID VALL\_TGMV  
AC P1557  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DE 01-JUN-1994 (Rel. 29, Last annotation update)  
DE ALI PROTEIN.  
GN ACI.  
OS Tomato golden mosaic virus (TGMV).  
OC Viruses: ssDNA viruses: Geminiviridae; Begomovirus.  
PR SEQUENCE FROM N.A.  
RA Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.  
RT \*Complete nucleotide sequence of the infectious cloned DNA components  
of tomato golden mosaic virus: potential coding regions and regulatory  
sequences.\*;  
RL EMBL J. 3:2197-2205(1984).  
CC -- SIMILARITY: BELONGS TO GEMINIVIRUSES ALI PROTEIN FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed, usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
EMBL: K02029; -- NOT ANNOTATED\_CDS.  
PIR: A04170; Q0CVL1  
DR INTERPRO: IPR001191;  
DR INTERPRO: IPR001301;  
DR PFAM: PF00799; GeminalALI; 1.  
DR PRINTS: PR00227; GEMCONTALI.  
DR PRINTS: PR00228; GEMCONTALVLI.  
DR ACDING: 223 230 ATP (POTENTIAL).  
FT NP\_BIND 223 230  
SQ SEQUENCE 352 AA; 40332 MW; C33C93BE9644B4A4 CRC64;

Query Match 97.0%; Score 350; DB 1; Length 352;  
Best Local Similarity 97.1%; Pred. No. 9.1e-32;  
Matches 68; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 TLWGFANVDSRGSGCQTNDAAEALNASSKEALQITREKIPKYLQFHNLNSL 60  
Db 111 TLWGFQVDCSRGSGCQTNDAAEALNASSKEALQITREKIPKYLQFHNLNSL 170

Oy 61 DRFDKTEP 70

Db 171 DRFDKTEP 180

RESULT 2

VALL\_PMYV STANDARD; PRT; 361 AA.

AC DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DE AL1 PROTEIN (40.4 KDA PROTEIN)  
 OS Potato yellow mosaic virus (isolate Venezuela).  
 SC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-91311403; PubMed-1856690;  
 RA Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.:  
 RT "The nucleotide sequence of the infectious cloned DNA components of  
 RT potato yellow mosaic virus", *Journal of Virology*, 72:1515-1520 (1991).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: D00340; BAA00782.1; -  
 CC PIR: J00364; Q0CVPT.  
 CC INTERPRO: IPR001131; -  
 CC PFAM: PF00759; GeminL.A1.1.  
 CC PRINTS: PR00227; GEMCOATL1.  
 CC PRINTS: PR00228; GEMCOATCLV1.  
 KW ATP-binding.  
 KW NP-BIND.  
 FT NP-BIND 222 229 ATP (POTENTIAL).  
 SQ SEQUENCE 361 AA: 40850 MW: 5627A33BF1264383 CRC64;

Query Match 69.3%; Score 250; DB 1; Length 361;  
 Best Local Similarity 68.1%; Pred. No. 1.3e-20;  
 Matches 47; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

QY 1 TLVGEAAVDGSRGCGQTSNDAAAFANASSKEALQITREKIPKYLQFHNLSNL 60  
 DB 110 TVMGQFQIDGSRGCGQTSNDAAAFANASSKEALQITREKIPKYLQFHNLSNL 169  
 QY 61 DRIFDKPTE 69  
 DB 170 DRIFDKAPE 178

RESULT 3  
 AC VAL1 CLVW STANDARD: PRT; 358 AA.  
 AC P14982;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DE AL1 PROTEIN (40.4 KDA PROTEIN)  
 OS Cassava latent virus (strain West Konyak 844).  
 SC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Stanley J., Gay M.R.;  
 RT "Nucleotide sequence of cassava latent virus DNA.",  
 RL *Nature* 301:260-262 (1983).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: J02057; A00111; NOT\_ANNOTATED\_CDS.  
 CC INTERPRO: IPR001131; -  
 CC PFAM: PF00759; GeminL.A1.1.  
 CC PRINTS: PR00227; GEMCOATL1.  
 CC PRINTS: PR00228; GEMCOATCLV1.  
 KW ATP-binding.  
 KW NP-BIND.  
 FT NP-BIND 220 227 ATP (POTENTIAL).  
 SQ SEQUENCE 358 AA: 40346 MW: ED173E753EE9D659 CRC64;

Query Match 65.7%; Score 237; DB 1; Length 358;  
 Best Local Similarity 61.4%; Pred. No. 3.5e-19;  
 Matches 43; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

QY 1 TLVGEAAVDGSRGCGQTSNDAAAFANASSKEALQITREKIPKYLQFHNLSNL 60  
 DB 109 TVMGQFQIDGSRGCGQTSNDAAAFANASSKEALQITREKIPKYLQFHNLSNL 166  
 QY 61 DRIFDKPTE 70  
 DB 169 DRIFQEPAP 178

RESULT 4  
 AC VAL1 CLVW STANDARD: PRT; 358 AA.  
 AC P14972;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DE AL1 PROTEIN (40.4 KDA PROTEIN)  
 OS Cassava latent virus (strain Nigerian).  
 SC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-90174930; PubMed-2308831;  
 RA Morris B., Coates L., Lowe S., Richardson K., Eddy P.;  
 RT "African cassava mosaic virus (Agerian strain).",  
 RL *Journal of Virology*, 64:197-198 (1990).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X17095; CAA34953.1; -  
 CC PIR: S07594; S07594.  
 CC INTERPRO: IPR001131; -  
 CC PFAM: PF001301; -  
 CC PRINTS: PR00227; GEMCOATL1.  
 CC PRINTS: PR00228; GEMCOATCLV1.  
 KW ATP-binding.  
 KW NP-BIND.  
 FT NP-BIND 220 227 ATP (POTENTIAL).  
 SQ SEQUENCE 358 AA: 40435 MW: 1D916B80CB2D5E2C CRC64;

Query Match 65.7%; Score 237; DB 1; Length 358;  
 Best Local Similarity 61.4%; Pred. No. 3.5e-19;  
 Matches 43; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

QY 1 TLVGEAAVDGSRGCGQTSNDAAAFANASSKEALQITREKIPKYLQFHNLSNL 60  
 DB 109 TVMGQFQIDGSRGCGQTSNDAAAFANASSKEALQITREKIPKYLQFHNLSNL 166

```

OY 61 DRIFTKTPP 70
DB 169 DRIFQPPAP 178

RESULT 5
VALL_TYLCU
ID VALL_TYLCU STANDARD; PRT; 362 AA.
AC P36279; 1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE ALL PROTEIN (C1 PROTEIN).
GN C1.
OS Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP SOURCE FROM N.A.
RX MEDLINE=9423446;
RA DRY I.B. Riden, J.E. Krake L.R., Mullineaux P.M., Hezalan M.A.;
RT "Nucleotide sequence and genome organization of tomato leaf curl
geminiivirus.";
RL J. Gen. Virol. 74:147-151(1993).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL; Z25751; CA81026.1; -.
DR PIR: S39211; S39211.
DR INTERPRO: IPR001191; -.
DR PFAM: PF00799; Gemini_AL1; 1.
DR PRINTS: PR00227; GEMCOATALL.
DR PRINTS: PR00228; GEMCOATCLVL.
KW ATP-binding. 221 228 ATP (POTENTIAL).
FT NP_BIND 221 228 ATP (POTENTIAL).
SQ SEQUENCE 362 AA; 41197 MW; 343271848704098 CRC64;

Query Match 63.7%; Score 230; DB 1; Length 362;
Best Local Similarity 54.1%; Pred. No. 2.1e-18;
Matches 46; Conservative 9; Mismatches 14; Indels 16; Gaps 1;

OY 1 TLWGEAANDGRSARGCOTSDAAEALNASSKEALQIREKIPKYLQFNHNSNL 60
DB 110 TLEWGEVQIDGRSARGCQSDAAQALNAGSKSEALNVLRELAPDYVLQFNHNSNL 169
OY 61 DRI-----FKTPE 69
DB 170 DRITPPLEVVSPFLSSFDNRPE 194

RESULT 6
VALL_TYLCU
ID VALL_TYLCU STANDARD; PRT; 359 AA.
AC P38609;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE ALL PROTEIN (C1 PROTEIN).
GN C1.
OS Tomato yellow leaf curl virus (strain Murcia) (TYLCV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP SOURCE FROM N.A.
RX MEDLINE=94256836; PubMed=8198442;

Noris E., Hidalgo E., Accotto G., Moriones E.;
"High similarity among the tomato yellow leaf curl virus isolates
from the west Mediterranean basin: the nucleotide sequence of an
infectious clone from Spain";
Nucleic Acids Res. 23:3920-3924, 1995.
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL; Z25751; CA81026.1; -.
DR PIR: S39211; S39211.
DR INTERPRO: IPR001191; -.
DR PFAM: PF00799; Gemini_AL1; 1.
DR PRINTS: PR00227; GEMCOATALL.
DR PRINTS: PR00228; GEMCOATCLVL.
KW ATP-binding. 221 228 ATP (POTENTIAL).
FT NP_BIND 221 228 ATP (POTENTIAL).
SQ SEQUENCE 359 AA; 41065 MW; 2D170A51EF80A3EC CRC64;

Query Match 51.8%; Score 223; DB 1; Length 359;
Best Local Similarity 43.4%; Pred. No. 1.2e-17;
Matches 41; Conservative 9; Mismatches 19; Indels 0; Gaps 0;

OY 2 TLWGEAANDGRSARGCOTSDAAEALNASSKEALQIREKIPKYLQFNHNSNL 61
DB 111 LEMWGTQIDGRSARGCQQTANDAYAKNAGSKSEALDVIRELAPROYLHFNHNSNL 170
OY 62 RIFQTPPEP 70
DB 171 RVFQVPPAP 179

RESULT 7
VALL_PHVU
ID VALL_PHVU STANDARD; PRT; 349 AA.
AC P34909;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE ALL PROTEIN.
GN ALL.
OS Pepper huasteco virus (PHV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP SOURCE FROM N.A.
RX MEDLINE=94015007; PubMed=8409944;
RA Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,
RA Rivera-Bustamante R.F.;
RT "Complete nucleotide sequence of pepper huasteco virus: analysis and
comparison with bipartite geminiviruses.";
J. Gen. Virol. 74:2225-2231(1993).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL; X70418; CA449856.1; -.
DR PIR: S31875; S31875.
DR JQ2300; JQ2300.
DR INTERPRO: IPR001191; -.

```



```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL: X15983; ; NOT_ANNOTATED_CDS.
DR YR: A35214; QOCVW1.
DR YR: A35214; QOCVW1.
DR INTERPRO: IPR001301; .
DR PFAM: PF00799; Gemini_L1.1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00228; GEMCOATCLV1.
DR ATP-binding. 221 228 ATP (POTENTIAL).
DR NP_BIND 335 AA; 40257 MW; 16A2CAN6A63251E95 CRC64;
SQ SEQUENCE 335 AA; 40257 MW; 16A2CAN6A63251E95 CRC64;

Query Match 59.0%; Score 213; DB 1; Length 355;
Best Local Similarity 57.1%; Pred. No. 1.6e-16; Mismatches 18; Indels 0; Gaps 0;
Matches 40; Conservative 12;

Oy 1 TLVNGEAVDGRSGCOTSDAAAEALNASSKEEALQIIRKTPKYLQFHNLSNL 60
Db 110 TENGFGQIDGRSGCQQTANGSYAKALMDVQSNLILKEEPQDFYQHNILNNA 169

Oy 61 DRIFDFTPEP 70
Db 170 ERIFAKPEP 179

RESULT 11
VAL1_TYLVY STANDARD; PRT; 357 AA.
AC P27259;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN (C1 PROTEIN).
OS Tomato yellow leaf curl virus (TYLCV).
CC This entry has been removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL: X15656; CAA33688.1; .
DR YR: D40778; QOCVCL.
DR INTERPRO: IPR001191; .
DR PFAM: PF00799; Gemini_L1.1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00228; GEMCOATCLV1.
DR ATP-binding. 219 226 ATP (POTENTIAL).
DR NP_BIND 357 AA; 40678 MW; 939ABG8B1AB3B2A7 CRC64;
SQ SEQUENCE 357 AA; 40678 MW; 939ABG8B1AB3B2A7 CRC64;

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL: X04144; ; NOT_ANNOTATED_CDS.
DR INTERPRO: IPR001191; .
DR PFAM: PF00799; Gemini_L1.1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00228; GEMCOATCLV1.
DR NP-binding. 222 229 ATP (POTENTIAL).
DR NP_BIND 358 AA; 40889 MW; 39A45FE3C0B9C333 CRC64;
SQ SEQUENCE 358 AA; 40889 MW; 39A45FE3C0B9C333 CRC64;

Query Match 59.2%; Score 210; DB 1; Length 358;
Best Local Similarity 54.3%; Pred. No. 3.6e-16;
Matches 38; Conservative 16; Mismatches 16; Indels 0; Gaps 0;

Oy 1 TLVNGEAVDGRSGCOTSDAAAEALNASSKEEALQIIRKTPKYLQFHNLSNL 60
Db 110 TENGFGQIDGRSGCQQTANGSYAKALMDVQSNLILKEEPQDFYQHNILNNA 169

Oy 61 DRIFDFTPEP 70
Db 170 OKITQPPPP 179

RESULT 13
VAL1_TMOV STANDARD; PRT; 361 AA.
ID VAL1_TMOV
DR YR: A35214; QOCVW1.
DR YR: A35214; QOCVW1.
DR 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN.
CC This entry has been removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----

```

```

Query Match 58.7%; Score 212; DB 1; Length 357;
Best Local Similarity 65.6%; Pred. No. 2.1e-16;
Matches 40; Conservative 11; Mismatches 10; Indels 0; Gaps 0;

Oy 4 MGEAVDGRSGCOTSDAAAEALNASSKEEALQIIRKTPKYLQFHNLSNLDRT 63
Db 111 FGVSDIGRSARGSGQSDANDAYAEALNASSKEEALNILKEAPKDYILOFHNLSNLDRI 170

Oy 64 F 64
Db 171 F 171

RESULT 12
VAL1_BCTV STANDARD; PRT; 358 AA.
ID VAL1_BCTV
AC P14591;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-APR-1992 (Rel. 23, Last annotation update)
DE AL1 PROTEIN (40.8 KDA PROTEIN).
OS Beet curly top virus (BCTV).
CC This entry has been removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL: X04144; ; NOT_ANNOTATED_CDS.
DR INTERPRO: IPR001191; .
DR PFAM: PF00799; Gemini_L1.1.
DR PRINTS: PR00227; GEMCOATL1.
DR PRINTS: PR00228; GEMCOATCLV1.
DR NP-binding. 222 229 ATP (POTENTIAL).
DR NP_BIND 358 AA; 40889 MW; 39A45FE3C0B9C333 CRC64;
SQ SEQUENCE 358 AA; 40889 MW; 39A45FE3C0B9C333 CRC64;

Query Match 58.2%; Score 210; DB 1; Length 358;
Best Local Similarity 54.3%; Pred. No. 3.6e-16;
Matches 38; Conservative 16; Mismatches 16; Indels 0; Gaps 0;

Oy 1 TLVNGEAVDGRSGCOTSDAAAEALNASSKEEALQIIRKTPKYLQFHNLSNL 60
Db 110 TENGFGQIDGRSGCQQTANGSYAKALMDVQSNLILKEEPQDFYQHNILNNA 169

Oy 61 DRIFDFTPEP 70
Db 170 OKITQPPPP 179

RESULT 13
VAL1_TMOV STANDARD; PRT; 361 AA.
ID VAL1_TMOV
DR YR: A35214; QOCVW1.
DR YR: A35214; QOCVW1.
DR 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN.
CC This entry has been removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----

```

OS Tomato mottle virus (isolate Florida) (TMov).  
 CC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.

RN [1] SEQUENCE FROM N.A.  
 RX MEDLINE-93107456; PubMed-1469361;  
 RA Abouzaid A.M., Polston J.E., Hiebert E.:  
 RT "The nucleotide sequence of tomato mottle virus, a new geminivirus  
 RL isolated from tomatoes in Florida.";  
 RL J. Gen. Virol. 73:3225-3229(1992).  
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC ENBL: L14460; AAC32414.1;  
 CC PIR: J01870; J01870;  
 DR INTERPRO: IPRO01191;  
 DR INTERPRO: IPRO01191;  
 DR PFAM: PF00799; Gemini\_ALI: 1.  
 DR PRINTS: PR00227; GEMCOATLVL.  
 KW ATP-binding; 322 220 ATP (BY SIMILARITY).  
 SQ SEQUENCE 361 AA; 40516 MW; 8138B65C8EAC6950 CRC64;

Query Match 58.2%; Score 210; DB 1; Length 361;  
 Best Local Similarity 54.3%; Pred. No. 3.6e-16;  
 Matches 38; Conservative 16; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLVMEADVRSRAGCGCSTNDAAALNAGSEALQIREKIPKYLQFYQFINLNSRL 60

DB 110 TLVMEADVRSRAGCGCSTNDAAALNAGSEALQIREKIPKYLQFYQFINLNSRL 169

QY 61 DRIFKPTPEP 70

DB 170 ERIFKAPPEP 179

RESULT 14

ID VALI\_SLCV STANDARD: PRT: 347 AA.

CC P29048;

DT 01-DEC-1992 (Rel. 24; Created)

DT 01-DEC-1992 (Rel. 24; Last sequence update)

DE ALL PROTEIN (Rel. 24; Last annotation update)

OS Squash leaf curl virus.

RN Viruses: ssDNA viruses; Geminiviridae; Begomovirus.

CC [1]

RX MEDLINE-51082449; PubMed-1984668;

RA Lazarowitz S.G., Tizard I.B.:

RT "Complete nucleotide sequence of the cloned genomic  
 RL component of the squash leaf curl geminivirus with a broad  
 RL host range phenotype.";  
 RL Virology 180:58-69(1991).

CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC ENBL: M3818; AAC32410.1; ALT-INIT.

DR

DR PIR: G36785; QOCVSL.  
 DR INTERPRO: IPRO01191;  
 DR PFAM: PF00799; Gemini\_ALI: 1.  
 DR PRINTS: PR00227; GEMCOATLVL.  
 KW ATP-binding; 218 225 ATP (POTENTIAL).  
 FT NP-BIND 347 AA; 39110 MW; AFDABDE122110E CRC64;  
 SQ SEQUENCE 347 AA; 39110 MW; AFDABDE122110E CRC64;

Query Match 32.4%; Score 117; DB 1; Length 347;  
 Best Local Similarity 31.3%; Pred. No. 1.1e-09;  
 Matches 25; Conservative 12; Mismatches 25; Indels 1; Gaps 1;

QY 5 GEADVRSRAGCGCSTNDAAALNAGSEALQIREKIPKYLQFYQFINLNSRLDRIF 64

DB 116 GQYKSG----GSKSKDQYINAVNAGSEALQIRKAGDKPTFTVYNHLLANVERLF 171

QY 65 DKTPPEP 70

DB 172 QKTPPEP 177

RESULT 15

ID LMA3\_HUMAN

CC Q16787; O13679; O13680;  
 CC 01-NOV-1997 (Rel. 35; Created)  
 CC 01-NOV-1997 (Rel. 35; Last sequence update)  
 CC 01-OCT-2000 (Rel. 40; Last annotation update)

DE LAMININ ALPHA-3 CHAIN PRECURSOR (EPILIGRIN 170 KDA SUBUNIT) (E170).

GN LAMA3.

OS Homo sapiens (Human).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

RN SEQUENCE FROM N.A.

CC TISSUE=KERATINOCYTES;

RX MEDLINE-94357926; PubMed-8077230;

RA Ryan M.C., Tizard R., Vandevanter D.R., Carter W.G.;

RT "Cloning of the LAMA3 gene encoding the alpha 3 chain of the adhesive  
 RL ligand epiligrin. Expression in wound repair.";  
 RL J. Biol. Chem. 269:22779-22787(1994).

RX SEQUENCE OF 1-1331 FROM N.A. (ISOFORMS A AND B).

RA MEDLINE-96163880; PubMed-8586427;

RA Vidal F., Baudoin C., Miquel C., Galliano M.-F., Cristiano A.M.;

RA Utito J., Ortonne J.-P., Meneguzzi G.;

RT "Cloning of the laminin alpha 3 chain gene (LAMA3) and identification  
 RL of a homozygous deletion in a patient with Herlitz junctional  
 RL epidermolysis bullosa.";  
 RL Genomics 30:273-280(1995).

CC -1- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ

CC IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF

CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT.

CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.

CC -1- FUNCTION: LAMININ-5 IS THOUGHT TO BE INVOLVED IN (1) CELL ADHESION

CC VIA INTEGRIN ALPHA-3/BETA-1 IN FOCAL ADHESION AND INTEGRIN ALPHA-

CC 3 IN CELL ADHESION AND CELL SPREAD, IN CELL SPREAD, IN CELL SPREAD,

CC PHOSPHORYLATION OF PP125-FAK AND P80, (3) DIFFERENTIATION OF

CC KERATINOCYTES.

CC -1- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE

CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND

CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE

CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.

CC THE ALPHA AND BETA CHAINS ARE IDENTICAL, WHILE THE GAMMA CHAIN IS

CC DIFFERENT AND POSSIBLY ALSO A COMPONENT OF LAMININ-6 (R-LAMININ)

CC AND LAMININ-7 (KS-LAMININ)

CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT

CC MEMBRANES (MAJOR COMPONENT).

CC -1- ALTERNATIVE PRODUCTS: THE TWO ISOFORMS A AND B DIFFER IN THEIR N-

CC TERMINUS. THE SEQUENCE SHOWN HERE IS THAT OF THE SMALLER VARIANT

CC

CC A.  
 CC -1- TISSUE SPECIFICITY: SKIN; RESPIRATORY, URINARY, AND DIGESTIVE  
 CC EPITHELIA AND IN OTHER SPECIALIZED TISSUES WITH PROMINENT  
 CC SECRETORY OR PROTECTIVE FUNCTIONS. EPITHELIAL BASEMENT MEMBRANE,  
 CC GLANDS, AND CYTOPLASMIC VESICLES. COLLAGEN TYPE I IS OBSERVED  
 CC DIFFERENTIAL AND FOCAL EXPRESSION OF THE ALPHA-3 CHAIN IS OBSERVED  
 CC IN THE CNS.  
 CC -1- INDUCTION: LAMININ-5 IS UP-REGULATED IN WOUND SITES OF HUMAN SKIN.  
 CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT  
 CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.  
 CC -1- DOMAIN: DOMAIN G IS GLOBULAR. XIS BULLOSA GRAVIS IS A BLISTERING  
 CC DISORDER IN SKIN THAT IS CHARACTERIZED BY A SEPARATION OF BASAL  
 CC CELLS FROM THE BASEMENT MEMBRANE DUE TO A DECREASED NUMBER OF  
 CC HEMIDESMOSOMES. LAMININ-5 IS MISSING FROM THE BASEMENT MEMBRANE OF  
 CC PATIENTS WITH THE GRAVIS FORM OF EPIDERMOLYSIS BULLOSA.  
 CC -1- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 2.5 LAMININ G-LIKE DOMAINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: L34155; AA559483.1; -;  
 CC INTERPRO: IPR001791; -;  
 CC EMBL: X85107; CA659428.1; -;  
 CC EMBL: X85108; CA659429.1; -;  
 CC HSP: P02468; IITL.  
 CC MW: 608085; -;  
 CC MI: 226700; -;  
 CC DR INTERPRO: IPR000561; -;  
 CC DR INTERPRO: IPR001791; -;  
 CC DR EMBL: X85107; CA659428.1; -;  
 CC DR PFAM: PF00053; laminin\_EGF\_2.  
 CC DR PFAM: PF00054; laminin\_G\_1.  
 CC DR PROSITE: PS00022; EGF\_1; 1.  
 CC DR PROSITE: PS01186; EGF\_2; 1.  
 CC DR PROSITE: PS01248; LAMININ\_TYPE\_EGF; 2.  
 CC KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;  
 CC MW Laminin EGF-like domain; Cell adhesion; Repeat; Signal;  
 CC SW A native sequence  
 CC FT SIGNAL 21 1713  
 CC FT CHAIN 21 1713 LAMININ ALPHA-3 CHAIN.  
 CC FT DOMAIN 46 201  
 CC FT DOMAIN 67 185 2.5 X LAMININ EGF-LIKE REPEATS.  
 CC FT DOMAIN 67 113 LAMININ EGF-LIKE 1.  
 CC FT DOMAIN 114 166 LAMININ EGF-LIKE 2.  
 CC FT DOMAIN 202 793 LAMININ EGF-LIKE 3. (INCOMPLETE).  
 CC FT DOMAIN 794 1713 LAMININ EGF-LIKE 4. (INCOMPLETE).  
 CC FT DOMAIN 794 1713 5 X LAMININ G-LIKE REPEATS (DOMAIN G).  
 CC FT DOMAIN 794 970 LAMININ G-LIKE 1.  
 CC FT DOMAIN 971 1139 LAMININ G-LIKE 2.  
 CC FT DOMAIN 1140 1353 LAMININ G-LIKE 3.  
 CC FT DOMAIN 1354 1529 LAMININ G-LIKE 4.  
 CC FT DOMAIN 1530 1713 LAMININ G-LIKE 5.  
 CC FT DOMAIN 396 548 COILED COIL (POTENTIAL).  
 CC FT DOMAIN 549 621 COILED COIL (POTENTIAL).  
 CC FT DOMAIN 594 621 COILED COIL (POTENTIAL).  
 CC FT DOMAIN 702 765 COILED COIL (POTENTIAL).  
 CC FT DOMAIN 1686 1713 COILED COIL (POTENTIAL).  
 CC FT DISULFID 67 76 BY SIMILARITY.  
 CC FT DISULFID 69 83 BY SIMILARITY.  
 CC FT DISULFID 86 95 BY SIMILARITY.  
 CC FT DISULFID 114 124 BY SIMILARITY.  
 CC FT DISULFID 116 135 BY SIMILARITY.  
 CC FT DISULFID 137 146 BY SIMILARITY.  
 CC FT DISULFID 149 164 BY SIMILARITY.  
 CC FT DISULFID 202 202 INTERCHAIN (PROBABLE).  
 CC FT DISULFID 205 205 INTERCHAIN (PROBABLE).

FT SITE 658 660  
 CC CARBOHYD 542 542 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 645 645 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 745 745 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 845 845 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 964 964 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 1108 1108 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 1131 1131 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 1325 1325 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 1477 1477 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 1667 1667 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC VANSOLIC 1 45  
 CC CONFLICT 5 5  
 CC CONFLICT 123 125  
 CC CONFLICT 481 481  
 CC CONFLICT 569 569  
 CC CONFLICT 969 969  
 CC CONFLICT 1052 1052  
 CC CONFLICT 1184 1184  
 CC SEQUENCE 1713 AA: 189304 MW: 45EA9BE107B6D03 CRC64:  
 19.0%; Score 68.5; DB 1; Length 1713;  
 Best Local Similarity 35.5%; Pred. No. 12;  
 Matches 22; Conservative 12; Mismatches 21; Indels 7; Gaps 3;  
 Oy 8 AVDGRARGGQTSNDAAAEALN--ASSKEALQ--TIPEKIPKYLQPHNINSDRIP 64  
 Db 552 AVDAATAVENILNATKAEDAAANRAASASALQTVIKEDLPK----AKTSSNDKLL 607  
 Oy 65 DK 66  
 Db 608 NE 609

Search completed: February 3, 2001, 02:24:29  
 Job time: 641 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.  
OM protein - protein search, using sw model  
Run on: February 3, 2001, 02:17:41 ; Search time 118.74 Seconds  
(without alignments)  
40.029 Million cell updates/sec

Title: US-09-289-346a-8  
Perfect score: 361  
Sequence: 1 TLVGEAAVDGSRGAGCQT.....PQFINLNSLDRFDKTEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues  
Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR66:\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	350	97.0	352	1 QOCV11	Al1 protein - toma
2	250	69.3	361	1 QOCVPT	Al1 protein - toma
3	237	65.7	358	2 S07594	hypothetical prote
4	230	63.7	362	1 JQ1887	Al1 protein - toma
5	223	61.8	359	2 S99211	gene Cl protein -
6	219	60.7	348	2 Q24796	replicase - pepper
7	215	59.8	359	2 S22553	Al1 protein - pep
8	216	59.8	359	2 S22553	Al1 protein - pep
9	213	59.0	351	2 J02327	Al1 protein - indi
10	213	59.0	355	1 QOCV41	AV1 protein - abut
11	212	58.7	357	1 QOCV1	Al1 protein - toma
12	210	58.2	358	1 J01870	Al1 protein - toma
13	210	58.2	359	2 S99235	gene Cl protein -
14	210	58.2	365	2 S28860	Al1 protein - beet
15	210	58.2	365	2 S28860	Al1 protein - associ
16	142	39.2	131	2 S45059	Al1 protein - squa
17	117	32.4	347	1 QOCV51	adhesive ligand ep
18	68.5	19.0	1713	2 A55347	FC gamma (Igc) rec
19	68	18.8	587	2 JCL1419	probable flagellar
20	66.5	18.4	201	2 A81360	probable peptidyl-
21	64.5	17.9	299	2 B71967	hypothetical prote
22	63.5	17.6	447	2 T12544	hypothetical prote
23	63.5	17.6	1646	2 J00422	voltage-dependent
24	63.5	17.6	1646	2 J00422	calcium channel al
25	63.5	17.6	2181	2 J00564	voltage-dependent
26	63.5	17.6	2181	2 A38198	calcium channel al
27	63.5	17.6	2203	2 T42742	Cl protein - tobac
28	62	17.2	295	2 D42452	probable phosphos
29	61.5	17.0	481	2 A70091	

30 61 16.9 1502 1 RGBYH1  
31 60 16.6 432 2 D75348  
32 60 16.6 447 2 S52437  
33 59.5 16.5 436 2 T24241  
34 59.5 16.5 356 2 T24241  
35 59.5 16.5 388 2 C59196  
36 59.5 16.5 2137 2 T05244  
37 59 16.3 316 2 C82085  
38 59 16.3 357 2 B71078  
39 58.5 16.2 297 1 MNVNRV  
40 58.5 16.2 335 2 J50164  
41 58.5 16.2 1084 2 T63465  
42 58.5 16.2 1084 2 T63465  
43 59.5 16.2 1070 1 A54600  
44 58.5 16.2 1265 2 T47626  
45 58.5 16.2 1743 2 T15893

RESULT 1  
QOCV11  
Al1 protein - tomato golden mosaic virus  
C:Species: tomato golden mosaic virus  
A:Note: host Nicotiana sp. (tobacco)  
C:Date: 28-Aug-1985 #sequence\_revision 28-Aug-1985 #text\_change 08-Apr-1994  
P:Hamilton, W.D.O.; Stehn, V.E.; Coutts, R.H.A.; Buck, K.W.  
EMBO J. 3, 2197-2205, 1984  
A:Title: Complete nucleotide sequence of the infectious cloned DNA components of toma  
A:Reference number: A04163  
A:Accession: A04170  
A:Molecule type: DNA  
A:Residues: 1-352 <HAMP>  
C:Comment: The genome consists of two circular, single-stranded DNA components, DNA A  
C:Genetics:  
A:Map position: segment A  
C:Superfamily: tomato golden mosaic virus Al1 protein

Query Match Local Similarity 97.0% Score 350 DB 1: Length 352;  
Matches 68; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Oy 1 TLVGEAAVDGSRGAGCQTSNDAAAEALNASSKEEALQIIRKIPKYLQFQHNLSNI. 60  
Db 111 TLVGEFVDGSRGAGCQTSNDAAAEALNASSKEEALQIIRKIPKYLQFQHNLSNI. 170  
Oy 61 DRIFDQTEP 70  
Db 171 DRIFDQTEP 180

RESULT 2  
QOCVPT  
Al1 protein - potato yellow mosaic virus (isolate Venezuela)  
C:Species: potato yellow mosaic virus  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Jun-2000  
C:Accession: J00364  
R:Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.  
J. Gen. Virol. 72, 1515-1520, 1991  
A:Title: The nucleotide sequence of the infectious cloned DNA components of potato ye  
A:Reference number: J00362; MUID:91311403  
A:Accession: J00364  
A:Status: translation not shown  
A:Residues: 1-361 <CCOU>  
A:Cross-references: GB:D00940; NID:g222458; PIDN:BAA00782.1; PID:g222459  
C:Genetics:  
A:Map position: segment A  
C:Superfamily: tomato golden mosaic virus Al1 protein

Db 110 TLEWGEFQIDGRSARGGQOSANDAYAQAALNTGSKSEALNVLRELAPKDYVLQFHNLSNL 169

```

Qy      61 DRI-----FDKTPE 69
      |||
      |||
db      170 DRIFTPLEVVSPELSSSEDRVPE 19

```

Db 110 TTWGLFQIDGRSARGGOOTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLQYHNLSCNL 169

QY 61 DRIFDKTPE 69  
 |||| | ||  
 DB 170 DRIFDKAPE 178

### RESULT 3

S07394  
hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)  
C:Date: 07-Sep-1990  
C:DB: 07-Sep-1990  
C:Text: 07-Sep-1990 #text\_change 20-Sep-1999  
C:Accession: S07594  
R:Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.  
Nucleic Acids Res. 18, 107-118, 1990  
A>Title: Nucleotide sequence of the infectious cloned DNA components of African cassava  
A:Reference number: S07590; MUD:90174930

A: Status; translation not shown  
A1: Molecule type: DNA  
A2: Residue type: A-358 <48>  
C: Sequences: EMBL:X17095; NID:959371; PID:CAN34953.1; PID:959376  
C1: Genetics  
C2: Superfamily: tomato golden mosaic virus All protein  
C3: Map position: segment DNA1

Query Match 65.7%; Score 237; DB 2; Length 358;  
Best Local Similarity 61.4%; Pred. No. 5.7e-19;  
Matches 43; Conservative 13; Mismatches 14; Indels

**OY**      1 TLVWGEAVDGRSARGGCOTSNDAAEALNASSKEEAQIREKIPKYLFQHHLNSNL 60  
              + : : :::::::::::::::::::: + : : ::::::::::  
**DG**     109 TWENGOFIDGRSGGOOSANDAKAINSGSKSEANVTRELVPRKFOTPHNT NSNL 168

Qy	61 DRIFDKTPEP 70
	:
Dh	160 DRIFGSDNAP 170

RESULT	4
JQ1887	
ALI protein - tomato yellow leaf curl virus (strain Australia)	

C;Species: tomato yellow leaf curl virus  
C;Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 07-May-1999  
C;Accession: JQ1887

J. Gen. Virol. 74, 147-151, 1993  
A; Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus  
A; Reference number: JQ1885; MUID: 93139778

A; Status: translation no

A; Residues: 1-362 <DRY>  
A; Cross-references: GR:S53251  
C; Superfamily: tomato golden mosaic virus All protein

Query Match	63.7%	Score 230;	DB 1;	Length 362;
Best Local Similarity	54.1%	Pred. No. 3.5e-18;		
Matches 46;	Conservative	9;	Mismatches 14;	Indels 16;
Gaps	1;			

Qy 1 TLVWGEAAVDGRSARGGQTSNDAAAEALNASSKEEALQIREKIDPKYLFQFHNLSNL 60

```
Qy 61 DRI-----FDKTP 69
    |||
    ||: ||
Db 170 DRIFTPLEVVVSPELSSSEDRP 194
```

RESULT 5  
S39211

gene C1 protein - tomato yellow leaf curl virus  
C:Species: tomato yellow leaf curl virus  
C.Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 20-Sep-1999  
C.Accession: S39211  
R.Norris, E.; Hidalgo, E.; Accotto, G.; Moriones, E.

A; Reference number: S39209

A: status: preliminary  
A: Molecule type: DNA  
A: Residues: 1-359 <NOR>  
A: Cross-references: EMBL:z25751; NID:g433655; PID:g433658  
C: Superfamily: tomato golden mosaic virus AL1 protein

Query Match 61.8%; Score 223; DB 2; Length 359;  
Best Local Similarity 59.4%; Pred. No. 2.1e-17;  
Matches 41; Conservative 9; Mismatches 19; Indels 0; Gaps 0;

Qy	2	LWGEAAVDGRSARGGCQTSDNAAABALNASSKEEALQIIREKIPEKYLQFHNLNSNLD	61
		_ _ _ _ _ : : : : : _ _ _ _ _ : : : : : _ _ _ _ _ : : : : : _ _ _ _ _ : : : : : _ _ _ _ _ : : : : : _ _ _ _ _ : : : : :	
Db	111	LEWGTQIDGRSARGGCQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD	170

QY 62 RIFDKTPEP 70

Db 171 RVFQVPPAP 179

## RESULT 6

JQ2300  
replicase - pepper huasteco virus (component A)  
N: Alternate names: ORF AL1 protein  
C: Species: pepper huasteco virus

C;Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 20-Sep-1999  
C;Accession: JQ2300  
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bucio  
74 0005 2003 1003

A:Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b  
A:Reference number: J02299; MUID:94015007

A:Residues: 1-349 <TOR>  
A:Cross-references: GB:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025  
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 60.7%; Score 219; DB 2; Length 349;  
Best Local Similarity 58.6%; Pred. No. 5.7e-17;  
Matches 41; Conservative 13; Mismatches 16; Indels

**Q7** 1 TLVWGAAVDGRSARGCQTSDNDAFAALNASSKEEALQIIPEKIPKYLQFPHNLNSL 60  
 : ||| :||||||| |::| :|||::| |||||::| |: : |||: ||  
**D6** 110 TVEWGEFIGRSARGCOOSANDYAKALNASAEPAALQIIEEPQHFFLFPHNTVSNA 169

Qy	61	DRFDKTEEP	70
	:		
Dh	170	NRTEOTPEEP	179

## RESULT 7

S31875  
 AL1 protein - pepper rizado amarillo virus  
 C:Species: pepper rizado amarillo virus



[illegible]

```

RESULT 12
QJ1870
A11 protein - tomato mottle virus (isolate Florida)
C:Species: tomato mottle virus
C:Date: 17-Feb-1994 #sequence-revision 17-Feb-1994 #text-change 07-May-1999
C:Accession: QJ1870
R:Adouid, A.M., Polston J.E., Hiebert, E.
A:Title: The nucleotide sequence of tomato mottle virus
A:Reference number: QJ1869; MUID:93107958
A:Accession: QJ1870
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-358 <ABO>
A:Cross-References: GB:I14460
A:Map position: segment A
C:Superfamily: tomato golden mosaic virus A11 protein

```

Query Match	58.2%	Score 210;	DB 1;	Length 358;
Best Local Similarity	54.3%	Pred. No. 6e-16;		
Matches 38; Conservative	16;	Mismatches 16;	Indels 0;	Gaps 0;

[illegible]

RESULT 13

S59235

gene C1 protein - tomato yellow leaf curl virus

C:Species: tomato yellow leaf curl virus

C:Date: 06-Jan-1995 #sequence-revision 06-Jan-1995 #text-change 08-Sep-1997

C:Accession: S39235

C:Author: E. Vaiera, A. Bascos, D. Accotto, G. Baccetti, and the EMBL data library

A:Description: A cloned DNA from a TYLCV isolate from Sicily showing low in

A:Reference number: S39233

A:Accession: S39235

A:Status: preliminary

A:Molecule type: ssDNA

A:Length: 1559

A:Cross-references: EMBL: Z28390; NID: g1041671; PID: g1334964

C:Superfamily: tomato golden mosaic virus A11 protein

Query Match 58.2%; Score 210; DB 2; Length 359;  
Best Local Similarity 55.1%; Pred. No. 6c-16;  
Matches 38; Conservative 12; Mismatches 19; Indels 0; Gaps 0;

```

QY 2 LWGEAAVDGRSARGGCOTSDAAAFALNASSKEALQIIREKTIPEKYLQFHNLSNLD 61
   || :||||||| ||:||||| ||||| ||||| :| :| :|||:||||
db 111 LEWGTFOIDGRSARGGQOTANDAYAKAINARSKSEALDVIKQLAPRDYVLHFHNLSNLD 170

```

```
QY 62 RIFDKTPEP 70
    ::| |
Db 171 KVFOVPPAP 179
```

RESULT 14

S28360

Al1 protein - beet curly top virus

C:Species: beet curly top virus

C:Date: 07-May-1993 #sequence:revision 07-May-1993 #text\_change 20-Sep-1999

A:Accession: S28360

A:Author: Callis, P.G.; Callis, R.J.; Pinner, M.S.

A:Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly top virus

A:Reference number: S28360

A:Accession: S28360

A:Status: translation not shown

A:Molecule type: DNA

A:Keywords: C:STAY>

A:Cross-references:GB:M24597; EMBL:X04144; NID:Q210678;

C:Species:tomato golden mosaic virus Al1 protein

PIDN:AAA42751.1; PID:Q210679

Query Match 58.28; Score 210; DB 2; Length 385;  
Best Local Similarity 54.38; Pred. No. 6.5e-16;  
Matches 38; Conservative 16; Mismatches 15; Indels

QY 1 TLWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPÉKYLQFQFHLNSNL 60  
I: III : IIIIIII III: : IIIII: : IIIII: : I: I III: I  
Db 137 TIEWGEFIDGRSARGGCQTANDSAKALNATSLDQALQILKXEQPKYFLOHINHLNNA 196

Qy	61	DRIFDKTPEP	70
		: : :	
Db	197	QKIFQRPPDP	206

RESULT 15

S59885

replication-associated protein C1 - tomato yellow leaf curl virus

C-Species: tomato yellow leaf curl virus

C-Date: 14-Jan-1996 #sequence\_revision 01-Mar-1996 #Text\_change 20-Sep-1999

C-Accession: S59885

R-Hong, Y.; Harrison, B.D.

A-Submitted to the EMBL Data Library, February 1995

A-Description: Nucleotide sequences from tomato leaf curl viruses from different geminiviruses.

A-Reference number: S59346

A-Accession: S19883

A-Accession: S59346

A-Molecule type: DNA

A-Residues: 1-360 <ON>

A-Cross-references: EMBL:Z48182; NID:g944838; PIDM:CAAG8229.1; PID:g974211

C-Superfamily: tomato golden mosaic virus A11 protein

Query Match	56.2%	Score 203;	DB 2;	Length 360;
Best Local Similarity	59.1%;	Pred. No. 3.7e-15;		
Matches 39: Conservative	12: Mismatches 15:	Indels 0: Gaps 0:		

**Qy**     **4**   WGEAAVDGSRSGCGQTSDNDAAAALNASSKEEALQIIREKIPEKYLFGFHNLNSLDR I   63  
      :  
      :  
**Dd**    113 FGVFIDGRSARGGOOSANDAYAEAINSGSKAALDILREKAPDEVLOFHNLNANLDRI 172

Qy	64	FDKTP	69
			:
Db	173	FTPSAE	178

Search completed: February 3, 2001, 02:17:41  
Job time: 2213 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: February 3, 2001, 02:19:48 ; Search time 109.09 Seconds  
(without alignments)  
11.523 Million cell updates/sec

Title: US-09-289-346a-8

Perfit score: 961

Sequence: 1 TLWGEAAVQGSARGCQT.....QFQHNLSNLDIFDKTPEP 70

Scoring table: BLOSSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/1/isa/5A.COMB.pep:\*

2: /cgn2\_6/ptodata/1/isa/5B.COMB.pep:\*

3: /cgn2\_6/ptodata/1/isa/6.COMB.pep:\*

4: /cgn2\_6/ptodata/1/isa/PCTUS.COMB.pep:\*

5: /cgn2\_6/ptodata/1/isa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result N	Score	Query Match %	Length	DB ID	Description
1	216	59.8	359	3	US-08-809-103B-2
2	216	59.8	359	3	US-08-809-103B-4
3	216	59.8	359	3	US-08-809-103B-6
4	216	59.8	359	3	US-08-809-103B-8
5	68.5	19.0	1713	3	US-08-809-982-24
6	68.5	17.6	2161	4	PCT/US94-10261A-24
7	68.5	17.6	2161	1	US-07-554-10261A-24
8	63.5	17.6	2161	1	US-08-455-543A-49
9	63.5	17.6	2161	1	US-08-455-543A-51
10	63.5	17.6	2161	2	US-08-223-305C-49
11	63.5	17.6	2161	2	US-08-223-305C-51
12	63.5	17.6	2161	2	US-08-311-363-2
13	60	16.6	446	2	US-08-672-814D-11
14	58	16.1	811	2	US-08-609-049A-22
15	58	16.1	811	2	US-08-609-049A-22
16	57.5	15.9	334	5	US20690-11
17	57.5	15.9	439	1	US-07-637-870-9
18	57.5	15.9	439	1	US-07-637-399-6
19	57.5	15.9	439	1	US-08-112-703-6
20	56	15.8	165	4	PCT-US95-10075-4
21	56	15.5	454	3	US-08-446-100-31
22	54.5	15.1	407	1	US-08-572-142A-13
23	54.5	15.1	407	1	US-08-572-142A-13
24	54.5	15.1	407	2	US-08-353-485-2
25	54.5	15.1	2628	2	US-08-570-311-14
26	54	15.0	454	3	US-08-446-100-30
27	54	15.0	602	2	US-08-419-652-6
28	54	15.0	771	1	US-07-923-976-6

Query Match 59.8% Score 216; DB 3; Length 359;  
Best Local Similarity 56.5%; Pred. No. 3,1e-20;  
Matches 39; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

Patent No. 5422248  
Sequence 4, Appl  
Sequence 8, Appl  
Sequence 1, Appl  
Sequence 6, Appl  
Sequence 2, Appl  
Sequence 26, Appl  
Sequence 27, Appl  
Sequence 28, Appl  
Sequence 17, Appl  
Sequence 18, Appl  
Sequence 23, Appl  
Patent No. 5231168  
Sequence 4, Appl  
Sequence 2, Appl

## ALIGNMENTS

RESULT 1  
US-08-809-103B-2  
; Sequence 2. Application US/0809103B  
; Patent No. 6133505  
; GENERAL INFORMATION:  
; APPLICANT: GRONBORN, BRUNO  
; INVENTOR: GRONBORN, BRUNO  
; TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS RESISTANT  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: YOUNG & THOMPSON  
; STREET: 745 South 23rd Street  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT ADDRESS: 10000 MATTHEW ROAD  
; APPLICATION NUMBER: US/08/809.103B  
; FILING DATE: 17-MAR-1997  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 94.11040  
; FILING DATE: 15-SEP-1994  
; PRIOR APPLICATION DATA: NO PCT/FR95/01192  
; APPLICATION NUMBER: NO  
; FILING DATE: 15-SEP-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: PATCH, Andrew J.  
; REGISTRATION NUMBER: 32,925  
; REFERENCE/DOCKET NUMBER: US94AL CNR TOM  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 521-2297  
; TELEFAX: (703) 521-0573  
; TELETYPE: 248423  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 359 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-809-103B-2

```

OY 2 LVWGEAVDGRSGCGQTSDAAAEALNASKEEALQIIRKIPKYLQFHNHNSLD 61
Db 111 LEMGTFOIDGRSGCGQTANDAYAKAINAGSKSOALDVIKELAPDYVLHFHNHNSLD 170
OY 62 RIFDKTPEP 70
Db 171 KVFQVPPAP 179

RESULT 2
US-08-809-103B-4
: Sequence 4, Application US/08809103B
: Patent No. 6133505
: GENERAL INFORMATION:
: APPLICANT: GRONENBORN, Bruno
: TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
: TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: YOUNG & THOMPSON
: STREET: 745 South 23rd Street
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/809,103B
: FILING DATE: 17-MAR-1997
: PRIORITY INFORMATION:
: PRIORITY NUMBER: 800
: APPLICATION NUMBER: FR 94.11040
: FILING DATE: 15-SEP-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WO PCT/FR95/01192
: FILING DATE: 15-SEP-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: PATCH, Andrew J.
: REGISTRATION NUMBER: 32,925
: REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 521-2297
: TELEFAX: (703) 685-0573
: INVOICING NUMBER: 4
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 359 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-809-103B-6

Query Match 59.8%; Score 216; DB 3; Length 359;
Best Local Similarity 56.5%; Pred. No. 3.1e-20;
Matches 39; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

OY 2 LVWGEAVDGRSGCGQTSDAAAEALNASKEEALQIIRKIPKYLQFHNHNSLD 61
Db 111 LEMGTFOIDGRSGCGQTANDAYAKAINAGSKSOALDVIKELAPDYVLHFHNHNSLD 170
OY 62 RIFDKTPEP 70
Db 171 KVFQVPPAP 179

RESULT 3
US-08-809-103B-6
: Sequence 4, Application US/08809103B
: Patent No. 6133505
: GENERAL INFORMATION:
: APPLICANT: GRONENBORN, Bruno
: TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
: TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: YOUNG & THOMPSON
: STREET: 745 South 23rd Street
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/809,103B
: FILING DATE: 17-MAR-1997
: PRIORITY INFORMATION:
: PRIORITY NUMBER: 800
: APPLICATION NUMBER: FR 94.11040
: FILING DATE: 15-SEP-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WO PCT/FR95/01192
: FILING DATE: 15-SEP-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: PATCH, Andrew J.
: REGISTRATION NUMBER: 32,925
: REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 521-2297
: TELEFAX: (703) 685-0573
: INVOICING NUMBER: 4
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 359 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-809-103B-4

Query Match 59.8%; Score 216; DB 3; Length 359;
Best Local Similarity 56.5%; Pred. No. 3.1e-20;
Matches 39; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

OY 2 LVWGEAVDGRSGCGQTSDAAAEALNASKEEALQIIRKIPKYLQFHNHNSLD 61
Db 111 LEMGTFOIDGRSGCGQTANDAYAKAINAGSKSOALDVIKELAPDYVLHFHNHNSLD 170
OY 62 RIFDKTPEP 70
Db 171 KVFQVPPAP 179

RESULT 4
US-08-809-103B-8
: Sequence 8, Application US/08809103B
: Patent No. 6133505
: GENERAL INFORMATION:
: APPLICANT: GRONENBORN, Bruno
: TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
: TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: YOUNG & THOMPSON
: STREET: 745 South 23rd Street
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/809,103B
: FILING DATE: 17-MAR-1997
: PRIORITY INFORMATION:
: PRIORITY NUMBER: 800
: APPLICATION NUMBER: FR 94.11040
: FILING DATE: 15-SEP-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WO PCT/FR95/01192
: FILING DATE: 15-SEP-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: PATCH, Andrew J.
: REGISTRATION NUMBER: 32,925
: REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 521-2297
: TELEFAX: (703) 685-0573
: INVOICING NUMBER: 4
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 359 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-809-103B-6

Query Match 59.8%; Score 216; DB 3; Length 359;
Best Local Similarity 56.5%; Pred. No. 3.1e-20;
Matches 39; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

OY 2 LVWGEAVDGRSGCGQTSDAAAEALNASKEEALQIIRKIPKYLQFHNHNSLD 61
Db 111 LEMGTFOIDGRSGCGQTANDAYAKAINAGSKSOALDVIKELAPDYVLHFHNHNSLD 170
OY 62 RIFDKTPEP 70
Db 171 KVFQVPPAP 179

```

```

: Sequence 6, Application US/08809103B
: Patent No. 6133505
: GENERAL INFORMATION:
: APPLICANT: GRONENBORN, Bruno
: TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
: TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: YOUNG & THOMPSON
: STREET: 745 South 23rd Street
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/809,103B
: FILING DATE: 17-MAR-1997
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: FR 94.11040
: FILING DATE: 15-SEP-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WO PCT/FR95/01192
: FILING DATE: 15-SEP-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: PATCH, Andrew J.
: REGISTRATION NUMBER: 32,925
: REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 521-2297
: TELEFAX: (703) 685-0573
: INVOICING NUMBER: 6:
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 359 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-809-103B-6

Query Match 59.8%; Score 216; DB 3; Length 359;
Best Local Similarity 56.5%; Pred. No. 3.1e-20;
Matches 39; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

OY 2 LVWGEAVDGRSGCGQTSDAAAEALNASKEEALQIIRKIPKYLQFHNHNSLD 61
Db 111 LEMGTFOIDGRSGCGQTANDAYAKAINAGSKSOALDVIKELAPDYVLHFHNHNSLD 170
OY 62 RIFDKTPEP 70
Db 171 KVFQVPPAP 179

RESULT 4
US-08-809-103B-8
: Sequence 8, Application US/08809103B
: Patent No. 6133505
: GENERAL INFORMATION:
: APPLICANT: GRONENBORN, Bruno
: TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
: TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: YOUNG & THOMPSON
: STREET: 745 South 23rd Street
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/809,103B
: FILING DATE: 17-MAR-1997
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: FR 94.11040
: FILING DATE: 15-SEP-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WO PCT/FR95/01192
: FILING DATE: 15-SEP-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: PATCH, Andrew J.
: REGISTRATION NUMBER: 32,925
: REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 521-2297
: TELEFAX: (703) 685-0573
: INVOICING NUMBER: 6:
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 359 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-809-103B-6

Query Match 59.8%; Score 216; DB 3; Length 359;
Best Local Similarity 56.5%; Pred. No. 3.1e-20;
Matches 39; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

OY 2 LVWGEAVDGRSGCGQTSDAAAEALNASKEEALQIIRKIPKYLQFHNHNSLD 61
Db 111 LEMGTFOIDGRSGCGQTANDAYAKAINAGSKSOALDVIKELAPDYVLHFHNHNSLD 170
OY 62 RIFDKTPEP 70
Db 171 KVFQVPPAP 179

```

```

? COUNTRY: U.S.A.
? ZIP: 2202
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/809,103B
? FILING DATE: 17-MAR-1997
? CLASSIFICATION: 800
? PRIORITY INFORMATION:
? PRIORITY NUMBER: FR 94,11040
? FILING DATE: 15-SEP-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: WO PCT/FR95/01192
? FILING DATE: 15-SEP-1995
? ATTORNEY/AGENT INFORMATION:
? NAME: PATCH, Andrew J.
? ADDRESS: 248425 EMBON
? REFERENCE/DOCKET NUMBER: US94AL CNR TOM
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (703) 521-2297
? TELEFAX: (703) 685-0573
? TELEX: 248425 EMBON
? INFORMATION FOR SEQ ID NO: 8:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1713 amino acids
? TYPE: amino acid
? TOPOLOGY: Linear
? MOLECULE TYPE: protein
? US-08-809-103B-8

Query Match 59.8%; Score 216; DB 3; Length 359;
Sequence Similarity 56.5%; Pctid No. 3; 5;
Matches 39; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

QY 2 LWNKEAAYDGRSGAGCGTSDNDAAEALN--ASSKEALO-IIREKIPEKYLQFHNLSNLD 61
DB 111 LWMGTQIDGRSGAGCGTSDNDAAEALN--ASSKEALO-IIREKIPEKYLQFHNLSNLD 170
QY 62 RIFDKTPEP 70
DB 171 KVFQVPEP 179

RESULT 5
? US-600-982-24
? Sequence 24, Application US/08600982
? Patent No. 6120991
? GENERAL INFORMATION:
? APPLICANT: Carter, William G.
? APPLICANT: GII, Susanna A.
? APPLICANT: Ryan, Maureen C.
? TITLE OF INVENTION: Epiligrin, an Epithelial Ligand for
? TITLE OF INVENTION: Integrins
? NUMBER OF SEQUENCES: 30
? CORRESPONDENCE ADDRESSES:
? ADDRESSEE: Christensen, O'Connor, Johnson, and Kindness
? STREET: 1420 Fifth Avenue
? CITY: Seattle
? STATE: WA
? COUNTRY: USA
? ZIP: 98101-8100
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: PCT/US94/10261A
? FILING DATE: 02-SEP-1994
? CLASSIFICATION:
? ATTORNEY/AGENT INFORMATION:
? NAME: Christensen, O'Connor, Johnson, and Kindness
? REGISTRATION NUMBER: 26,997
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (206) 682-8100
? TELEFAX: (206) 224-0779
? INFORMATION FOR SEQ ID NO: 24:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1713 amino acids
? TYPE: amino acid
? TOPOLOGY: Linear
? MOLECULE TYPE: protein
? DESCRIPTION: E170 protein as translated from sequence of
? APPLICATION: FIGURES 15A-15F, and as shown also in FIGURES 19A-19F
? FILING DATE: 02-SEP-1994

```

```

? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Shelton, Dennis K.
? REGISTRATION NUMBER: 26,997
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (206) 682-8100
? TELEFAX: (206) 224-0779
? INFORMATION FOR SEQ ID NO: 24:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1713 amino acids
? TYPE: amino acid
? TOPOLOGY: Linear
? MOLECULE TYPE: protein
? DESCRIPTION: E170 protein as translated from sequence
? DESCRIPTION: OF FIGURES 15A-15F, and as shown also in FIGURES
? DESCRIPTION: 19A-19F
? US-08-600-982-24

Query Match 19.0%; Score 68.5; DB 3; Length 1713;
Sequence Similarity 35.3%; Pctid No. 3; 5;
Matches 22; Conservative 12; Mismatches 21; Indels 7; Gaps 3;

QY 8 AVDCRSARGCGTSDNDAAEALN--ASSKEALO-IIREKIPEKYLQFHNLSNLDRIIF 64
DB 552 AVDAATAVETNLNKAEDAAANRAASASALQTVIKEDLPK----AKTSSNSDKLL 607
QY 65 DK 66
DB 608 NE 609

RESULT 6
? PCT-US94-10261A-24
? Sequence 24, Application PCT/US9410261A
? GENERAL INFORMATION:
? APPLICANT: Carter, William G.
? APPLICANT: GII, Susanna A.
? APPLICANT: Ryan, Maureen C.
? TITLE OF INVENTION: Epiligrin, an Epithelial Ligand for
? TITLE OF INVENTION: Integrins
? NUMBER OF SEQUENCES: 30
? CORRESPONDENCE ADDRESSES:
? ADDRESSEE: Christensen, O'Connor, Johnson, and Kindness
? STREET: 1420 Fifth Avenue
? CITY: Seattle
? STATE: WA
? COUNTRY: USA
? ZIP: 98101-8100
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: PCT/US94/10261A
? FILING DATE: 02-SEP-1994
? CLASSIFICATION:
? ATTORNEY/AGENT INFORMATION:
? NAME: Christensen, O'Connor, Johnson, and Kindness
? REGISTRATION NUMBER: 26,997
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (206) 682-8100
? TELEFAX: (206) 224-0779
? INFORMATION FOR SEQ ID NO: 24:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1713 amino acids
? TYPE: amino acid
? TOPOLOGY: Linear
? MOLECULE TYPE: protein
? DESCRIPTION: E170 protein as translated from sequence of
? APPLICATION: FIGURES 15A-15F, and as shown also in FIGURES 19A-19F
? FILING DATE: 02-SEP-1994

```

Query Match 17.6%; Score 63.5; DB 1; Length 2161;  
Best Local Similarity 28.4%; Pred. No. 21;  
Matches 21; Conservative 5; Mismatches 19; Indels 29; Gaps 1;

RESULT 8  
US-08-455-543A-49





REFERENCE/DOCKET NUMBER: 52516 (P519739)  
TELECOMMUNICATION INFORMATION:  
FILING DATE: 04-APR-1989  
APPLICATION NUMBER: US 07/176,899  
INVENTOR: (619)238-0062 49:  
ATTORNEY/AGENT INFORMATION:  
NAME: Seidman, Stephanie L.  
LENGTH: 2161 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: Internal  
US-08-223-303C-49

Query Match 17.6% Score 63.5; DB 2; Length 2161;  
Best Local Similarity 28.4% Pred. No. 21;  
Matches 21; Conservative 5; Mismatches 19; Indels 29; Gaps 1;  
QY 4 NCEAAYDGRSGAGCGQTS-----NDAAAEALNASK 34  
DB 707 NNAVMTDGINAVGPPSSGMIVCFITLFCGNYLLINVLATAVDNLADESINTAQK 766

QY 35 EEAQLQIREKPEK 48  
DB 767 EEAEEKERKXK 780

RESULT 11  
US-08-223-303C-51  
Sequence 51, Application US/08223303C  
Patent No. 5851824  
GENERAL INFORMATION:  
APPLICANT: Harpold, Michael  
APPLICANT: Ellis, Steven  
APPLICANT: Williams, Mark  
APPLICANT: Feldman, Daniel  
APPLICANT: McCue, Ann  
APPLICANT: Brenner, Robert  
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND  
METHODS  
NUMBER OF SEQUENCES: 57  
CORRESPONDENCE ADDRESS: 32  
ADDRESS: 1660 Union Street  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-2926  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/223,303C  
FILING DATE: April 4, 1994  
PRIOR APPLICATION DATA:  
APPLICATION DATA: 07/868,354  
FILING DATE: April 10, 1992  
PRIOR APPLICATION NUMBER: US 07/745,206  
FILING DATE: 15-AUG-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/620,250  
FILING DATE: 20-FEB-1990  
PRIOR APPLICATION NUMBER: US 07/482,384  
FILING DATE: 20-FEB-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/603,751  
FILING DATE: 04-APR-1989  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US89/01408  
FILING DATE: 04-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/176,899  
INVENTOR: (619)238-0062 49:  
ATTORNEY/AGENT INFORMATION:  
NAME: Seidman, Stephanie L.  
REGISTRATION NUMBER: 33,779  
REFERENCE/DOCKET NUMBER: 52516 (P519739)  
TELECOMMUNICATION INFORMATION:  
FILING DATE: 04-APR-1989  
APPLICATION NUMBER: US 07/176,899  
INVENTOR: (619)238-0062 51:  
ATTORNEY/AGENT INFORMATION:  
NAME: Seidman, Stephanie L.  
LENGTH: 2161 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: Internal  
US-08-223-303C-51

Query Match 17.6% Score 63.5; DB 2; Length 2161;  
Best Local Similarity 28.4% Pred. No. 21;  
Matches 21; Conservative 5; Mismatches 19; Indels 29; Gaps 1;  
QY 4 NCEAAYDGRSGAGCGQTS-----NDAAAEALNASK 34  
DB 707 NNAVMTDGINAVGPPSSGMIVCFITLFCGNYLLINVLATAVDNLADESINTAQK 766

QY 35 EEAQLQIREKPEK 48  
DB 767 EEAEEKERKXK 780

RESULT 12  
US-08-311-363-2  
Sequence 2, Application US/08311363  
Patent No. 5876958  
GENERAL INFORMATION:  
APPLICANT: Harpold, Michael  
APPLICANT: Ellis, Steven  
APPLICANT: Williams, Mark  
APPLICANT: Feldman, Daniel  
APPLICANT: McCue, Ann  
APPLICANT: Brenner, Robert  
TITLE OF INVENTION: Human Calcium Channel Compositions and  
Methods  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS: 32  
ADDRESS: 1660 Union Street  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-2926  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/311,363  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/745,206  
FILING DATE: 15-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Seidman, Stephanie L.  
REGISTRATION NUMBER: 33,779  
REFERENCE/DOCKET NUMBER: 6362-51506  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (619)238-0999  
 TELEFAX: (619)238-0062  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-311-363-2

Query Match 17.6%; Score 63.5; DB 2; Length 2161;  
 Best Local Similarity 28.4%; Pred. No. 21;  
 Matches 21; Conservative 5; Mismatches 19; Indels 29; Gaps 1;  
 QY 4 WEEAADVGSARGGQTS-----NDAARALNASSK 34  
 DB 707 WNAVMDIGMAYGSSSGMVICVIFILPGNYILLNVFLIAVINDLAESLNTAQK 766

QY 35 EEAOLIIPEKX 48  
 DB 767 EEAEEKKIKARK 780

RESULT 13  
 US-08-672-814D-11  
 Sequence 11, Application US/08672814D  
 Patent No. 5948664  
 GENERAL INFORMATION:  
 APPLICANT: Leung, David W.  
 TITLE OF INVENTION: MAMMALIAN  
 NUMBER OF SEQUENCES: 13  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Cell Therapeutics, Inc.  
 CITY: Seattle  
 STATE: Washington  
 COUNTRY: U.S.A.  
 ZIP: 98119  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5" disk, 1.44MB, double side, high density  
 CONTAINER: One (486 microprocessor)  
 OPERATING SYSTEM: MS-DOS Version 6.1, Windows NT  
 SOFTWARE: WORD 6.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/672,814D  
 FILING DATE: 28-Jun-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Falciszewski, Stephen  
 TELEPHONE: (206)284-6131  
 REFERENCE/DOCKET NUMBER: 2307K-06370003  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206)284-7100  
 TELEFAX: (206)284-6206  
 INFORMATION FOR SEQ ID NO: 11:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 446  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 HYPOTHETICAL: no  
 ANTI-SENSE: no  
 FRAGMENT TYPE:  
 ORIGINAL SOURCE:  
 ORGANISM: Drosophila  
 DEVELOPMENTAL STAGE:  
 INDIVIDUAL ISOLATE:  
 HAPLOTYPE:  
 TISSUE TYPE:  
 CELL TYPE:

CELL LINE:  
 ORGANELLE:  
 FEATURE:  
 NAME/KEY: CDP-diacylglycerol synthase (CDS)  
 LOCATION:  
 IDENTIFICATION METHOD:  
 OTHER INFORMATION:  
 US-08-672-814D-11

Query Match 16.6%; Score 60; DB 2; Length 446;  
 Best Local Similarity 34.0%; Pred. No. 7.8;  
 Matches 17; Conservative 9; Mismatches 22; Indels 2;  
 Gaps 1;  
 QY 20 TSDAAAEALNASSKEEAQIIPEKYLQFQINLSNLDRIPTKE 69  
 DB 20 SGSDAANKNSAADSDHVDSEETPEEKFDV--ELKNLPQCTDKTPE 67

RESULT 14  
 US-08-609-049A-22  
 Sequence 22, Application US/08609049A  
 Patent No. 5948664  
 GENERAL INFORMATION:  
 APPLICANT: Williams, Lewis T.  
 APPLICANT: Molz, Lisa  
 APPLICANT: Chen, Yen-Wen  
 TITLE OF INVENTION: NOVEL  
 NUMBER OF SEQUENCES: 2  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Townsend and Townsend and Crew LLP  
 STREET: Two Embarcadero Center, 8th Floor  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94111-3834  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/609,049A  
 FILING DATE: 29-Feb-1996  
 CLASSIFICATION: I35  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Dow, Karen B.  
 REGISTRATION NUMBER: 29,684  
 REFERENCE/DOCKET NUMBER: 2307K-06370005  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-326-2400  
 TELEFAX: 415-326-2442  
 INFORMATION FOR SEQ ID NO: 22:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 171 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-609-049A-22

Query Match 16.1%; Score 58; DB 2; Length 171;  
 Best Local Similarity 24.7%; Pred. No. 4.1;  
 Matches 22; Conservative 11; Mismatches 20; Indels 36; Gaps 3;  
 QY 2 LVWGEAAVDGSRAGGC-----QTSNDAAEALNASSKEA 37  
 DB 24 LIRKEAGLDIRMLPGCLATGDSGLIEWYSITETADQLNSNVAAAAA--AFRDL 81  
 QY 38 LQIIREKIPKYLQFQINLSNLDRIPTKE 66  
 DB 82 LANLKE-----YNSGDDLRAIEE 100



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:15:29 ; Search time 144.12 seconds  
(without alignments)  
16.608 Million cell updates/sec

Title: US-09-289-346a-8

Perfect score: 361

Sequence: 1 TLVWGAAVDGSRGCGQT.....FOFHNLNSLDRIEDKTEPP 70

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 segs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_36\_\*  
1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Score	Query Match	Length	DB ID	Description
1	216	59.8	353	18 W34338	Bean golden mosaic
2	216	59.8	353	18 W34332	Bean golden mosaic
3	216	59.8	353	18 W34334	Bean golden mosaic
4	216	59.8	353	18 W34335	Bean golden mosaic
5	216	59.8	353	18 W34336	Bean golden mosaic
6	216	59.8	359	17 R88870	Sardinian tomato Y
7	216	59.8	359	17 R88871	Sardinian tomato Y
8	216	59.8	359	17 R88872	Sardinian tomato Y
9	214	59.3	353	6 P70407	ORF 4 gene product Y
10	214	59.3	361	18 W34336	Tomato mottle virus
11	214	59.3	361	18 W34324	Tomato mottle virus
12	214	59.3	361	18 W34325	Tomato mottle virus

13	214	59.3	361	18 W34326	Tomato mottle virus
14	212	58.7	357	18 W34329	Tomato yellow leaf
15	212	58.7	357	18 W34330	Tomato yellow leaf
16	212	58.7	357	18 W34331	Tomato yellow leaf
17	202	56.5	357	18 W34332	Tomato yellow leaf
18	202	56.5	357	18 W34333	Tomato yellow leaf
19	202	56.0	362	19 W56495	Tobacco leaf curl
20	68.5	19.0	1713	16 W70148	Deduced sequence o
21	66.5	18.4	512	19 W68473	HIV-1 strain YBF30
22	65	18.0	252	21 Y92317	Measlesvirus Rep C1
23	65	18.0	335	21 Y92318	Measlesvirus Rep C1
24	64.5	17.9	299	18 W54550	H. pylori ORF 02ae
25	63.5	17.6	447	21 W65748	Human secreted pro
26	63.5	17.6	447	21 W65749	Human secreted pro
27	63.5	17.6	447	21 W50947	Human adult aorta
28	63.5	17.6	456	21 W87190	Human secreted pro
29	63.5	17.6	619	13 R27651	Human calcium chan
30	63.5	17.6	2161	14 R35545	Sequence of the al
31	63.5	17.6	2161	16 R71001	Human neuronal cal
32	63.5	17.6	2161	16 R71002	Human neuronal cal
33	63.5	17.6	2161	19 W63132	Human calcium chan
34	63.5	17.6	2161	19 W63139	Human calcium chan
35	60.5	16.8	204	20 W89614	Protein encoded by
36	60.5	16.8	299	20 W89629	Protein encoded by
37	60.5	16.8	299	20 W89649	Antigen from clust
38	59.5	16.5	299	19 W98322	H. pylori GHPO 136
39	59.5	16.5	299	19 W73034	Helicobacter pylori
40	58.5	16.1	844	18 P00773	Protein M1 encoded
41	58.5	16.1	844	18 P00773	Protein M1 encoded
42	57.5	15.9	334	11 R03211	Amio acid sequenc
43	57	15.8	165	17 R91063	Turkey prepro-VIP.
44	56.5	15.7	249	20 W89622	Protein encoded by
45	56.5	15.7	269	20 W89650	Antigen 1 from clu

#### ALIGNMENTS

RESULT 1  
W34338 ID W34338 standard; Protein: 353 AA.  
XX AC W34338;  
XX AC W34338;  
XX DT 27-APR-1998 (first entry)  
XX DE Bean golden mosaic geminivirus C1 protein.  
XX GN Geminivirus; BGWV; C1 gene; transdominant mutation;  
XX KN transgenic plant; disease resistance.  
XX XX Bean golden mosaic virus type II isolate Guatemala.  
XX W09739110-A1.  
XX DT 23-OCT-1997.  
XX PF 15-APR-1997; 97WO-US06300.  
XX XX 16-APR-1996; 96US-0015517.  
XX XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.  
XX (WISC) WISCONSIN ALUMNI RES FOUND.  
XX Ahlquist PG, Hanson SF, Lau HT, Maxwell DP, Stout JT;  
WPI: 1997-526447/48.  
N-PSDB; 193314.

Transgenic plants expressing geminivirus AC1 and C1 wild-type and mutant genes - have increased resistance to geminivirus infection  
e.g. tomato mottle virus, tomato yellow leaf curl virus or bean golden mosaic geminivirus

xx Example 5; Page 100-102; 132pp; English.

xx This sequence comprises the bean golden mosaic virus (BGMV) CI protein that is encoded for in the BGMV genome. It involves production of transgenic plants containing DNA comprising CI or AC1 wild-type or mutant sequences that negatively interfere in trans with geminiviral replication during infection. Such transgenic plants are resistant to viral infection. The AC1/CI genes are especially from BGW, tomato mottle virus or tomato yellow leaf curl virus (see T93282-93) and encode polypeptides (see W34324-35) that have mutations in the highly conserved DNA-nicking domain CC and/or the NTP-binding domains.

xx Sequence 353 AA:

SQ

Query Match 59.8%; Score 216; DB 18; Length 353;  
Best Local Similarity 58.8%; Pred. No. 3e-19;  
Matches 41; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

OY 1 TLVKGEEAVDGRSGGQTSNDAAALNASSKEALQIRKIPKYLQFHNLSNL 60  
Db 110 TLVKGEEAVDGRSGGQTSNDAAALNASSKEALQIRKIPKYLQFHNLSNL 60  
OY 61 DRIFDKTPEP 70  
Db 170 DRIFDKTPEP 179

RESULT 2

ID W34332 standard; Protein; 353 AA.

XX W34332;

XX W34332;

DT 27-APR-1998 (first entry)

DE Bean golden mosaic geminivirus CI BGA190 mutant.

XX Geminivirus; BGMV; CI gene; transdominant mutation;  
KW transgenic plant; disease resistance.

XX Bean golden mosaic virus type II isolate Guatemala.

OS W09739110-A1.

PN W09739110-A1.

XX W09739110-A1.

PD 23-OCT-1997.

XX 15-APR-1997; 97WNO-US06300.

XX 16-APR-1996; 96US-0015517.

PR (SEMI-) SEMINIS VEGETABLE SEEDS INC.

PA (WISC-) WISCONSIN ALUMNI RES FOUND.

XX Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;  
PI WPI; 1997-526447/48.  
DR N-PSDB; T93290.

XX Transgenic plants expressing geminivirus AC1 and CI wild-type and mutant genes - have increased resistance to geminivirus infection e.g. tomato mottle virus, tomato yellow leaf curl virus or bean golden mosaic geminivirus

XX Example 5; Page 103-105; 132pp; English.

xx This protein comprises a control mutant of the bean golden mosaic virus (BGMV) CI protein (see W34338) that is required for replication. It is encoded by mutated CI open reading frame CC BGAC130 (see T93290). The invention involves production of transgenic plants containing DNA comprising geminivirus CI or AC1

CC wild-type or mutant sequences that negatively interfere in trans with geminiviral replication during infection. Such transgenic plants are resistant to viral infection. The AC1/CI genes are especially from BGW, tomato mottle virus or tomato yellow leaf curl virus (see T93282-93) and encode polypeptides (see W34324-35) that have mutations in the highly conserved DNA-nicking and/or the NTP-binding domains.

xx Sequence 353 AA:

SQ

Query Match 59.8%; Score 216; DB 18; Length 353;  
Best Local Similarity 58.8%; Pred. No. 3e-19;  
Matches 41; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

OY 1 TLVKGEEAVDGRSGGQTSNDAAALNASSKEALQIRKIPKYLQFHNLSNL 60  
Db 110 TLVKGEEAVDGRSGGQTSNDAAALNASSKEALQIRKIPKYLQFHNLSNL 60  
OY 61 DRIFDKTPEP 70  
Db 170 DRIFDKTPEP 179

RESULT 3

ID W34333 standard; Protein; 353 AA.

XX W34333;

XX W34333;

DT 27-APR-1998 (first entry)

DE Bean golden mosaic geminivirus CI BGA221 mutant.

XX Geminivirus; BGMV; CI gene; transdominant mutation;  
KW transgenic plant; disease resistance.

XX Bean golden mosaic virus type II isolate Guatemala.

OS W09739110-A1.

PN W09739110-A1.

XX W09739110-A1.

PD 23-OCT-1997.

XX 15-APR-1997; 97WNO-US06300.

XX 16-APR-1996; 96US-0015517.

PR (SEMI-) SEMINIS VEGETABLE SEEDS INC.

PA (WISC-) WISCONSIN ALUMNI RES FOUND.

XX Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;  
PI WPI; 1997-526447/48.  
DR N-PSDB; T93291.

XX Transgenic plants expressing geminivirus AC1 and CI wild-type and mutant genes - have increased resistance to geminivirus infection e.g. tomato mottle virus, tomato yellow leaf curl virus or bean golden mosaic geminivirus

XX Example 5; Page 107-109; 132pp; English.

xx This protein comprises a transdominant lethal mutant of the bean golden mosaic virus (BGMV) CI protein (see W34338) that is required for replication. It is encoded by mutated CI open reading frame CC BGAC221 (see T93291) and carries a mutation in the NTP-binding domain of the protein. The invention involves production of transgenic plants comprising geminivirus CI or AC1 wild-type or mutant sequences that negatively interfere in trans with geminiviral replication during infection. Such transgenic plants are resistant to viral infection. The AC1/CI genes are especially from BGW, tomato mottle virus or tomato yellow leaf curl virus (see T93282-93) and encode polypeptides (see W34324-35) that have mutations in the

CC highly conserved DNA-nicking and/or the NTP-binding domains.

XX Sequence 353 AA;

Query Match 59.8%; Score 216; DB 18; Length 353;  
Best Local Similarity 58.6%; Pred. No. 3e-19;  
Matches 41; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

Qy 1 TLVWGEAAVGRSGAGCGCOTSDAAAEALNASSKEEALQIIREIKPKYLFQPHLNSNL 60  
Db 1 TLVWGEAAVGRSGAGCGCOTSDAAAEALNASSKEEALQIIREIKPKYLFQPHLNSNL 169  
Qy 61 DRIFDKTPEP 70  
Db 170 erifvkpep 179

RESULT 4  
W34334  
ID W34334 standard; Protein: 353 AA.

AC W34334;  
XX Bean golden mosaic geminivirus Cl BG4228 mutant.  
XX 27-APR-1998 (first entry)  
XX Bean golden mosaic geminivirus Cl BG4228 mutant.  
XX Geminivirus; BGW; Cl gene; transdominant mutation;  
XX transgenic plant; disease resistance.  
XX Bean golden mosaic virus type II isolate Guatemala.

XX W09739110-AL.  
XX 23-OCT-1997.  
XX 15-APR-1997; 97WO-US06300.  
XX 16-APR-1996; 96US-0015517.

(SMT-) SEMINIS VEGETABLE SEEDS INC.  
(WISC) WISCONSIN ALUMNI RES FOUND.

PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;  
XX WPI: 1997-526447/48.  
XX N-PSDB: T93292.

XX Transgenic plants expressing geminivirus AC1 and CI wild-type and  
XX mutant genes have increased resistance to geminivirus infection  
XX e.g. tomato mottle virus, tomato yellow leaf curl virus or bean  
XX golden mosaic geminivirus

XX Example 5; Page 111-112; 132pp; English.

XX This protein comprises a transdominant lethal mutant of the bean  
XX golden mosaic virus (BGW) CI protein (see W34334) that is required  
XX for replication. It is encoded by mutated CI open reading frame  
XX BGAC263 (see T93292) and carries a mutation in the NTP-binding  
XX domain. The invention involves production of transgenic plants  
XX containing DNA comprising geminivirus CI or AC1 wild-type or mutant  
XX sequences that negatively interfere in trans with geminiviral  
XX replication during infection. Such transgenic plants are resistant  
XX to viral infection. The AC1/CI genes are especially from BGW,  
XX tomato mottle virus or tomato yellow leaf curl virus (see T93292-93)  
XX and encode polypeptides (see W34334-35) that have mutations in the  
XX highly conserved DNA-nicking and/or the NTP-binding domains.

XX Sequence 353 AA;

Query Match 59.8%; Score 216; DB 18; Length 353;

Best Local Similarity 58.6%; Pred. No. 3e-19;  
Matches 41; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

Qy 1 TLVWGEAAVGRSGAGCGCOTSDAAAEALNASSKEEALQIIREIKPKYLFQPHLNSNL 60  
Db 1 TLVWGEAAVGRSGAGCGCOTSDAAAEALNASSKEEALQIIREIKPKYLFQPHLNSNL 169  
Qy 61 DRIFDKTPEP 70  
Db 170 erifvkpep 179

RESULT 5  
W34335  
ID W34335 standard; Protein: 353 AA.

AC W34335;  
XX 27-APR-1998 (first entry)  
XX Bean golden mosaic geminivirus Cl BG4262 mutant.  
XX Geminivirus; BGW; Cl gene; transdominant mutation;  
XX transgenic plant; disease resistance.  
XX Bean golden mosaic virus type II isolate Guatemala.

XX W09739110-AL.  
XX 23-OCT-1997.  
XX 15-APR-1997; 97WO-US06300.  
XX 16-APR-1996; 96US-0015517.

(SMT-) SEMINIS VEGETABLE SEEDS INC.  
(WISC) WISCONSIN ALUMNI RES FOUND.

PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;  
XX WPI: 1997-526447/48.  
XX N-PSDB: T93293.

XX Transgenic plants expressing geminivirus AC1 and CI wild-type and  
XX mutant genes have increased resistance to geminivirus infection  
XX e.g. tomato mottle virus, tomato yellow leaf curl virus or bean  
XX golden mosaic geminivirus

XX Example 5; Page 115-116; 132pp; English.

XX This protein comprises a transdominant lethal mutant of the bean  
XX golden mosaic virus (BGW) CI protein (see W34335) that is required  
XX for replication. It is encoded by mutated CI open reading frame  
XX BGAC263 (see T93293) and carries a mutation in the NTP-binding  
XX domain. The invention involves production of transgenic plants  
XX containing DNA comprising geminivirus CI or AC1 wild-type or mutant  
XX sequences that negatively interfere in trans with geminiviral  
XX replication during infection. Such transgenic plants are resistant  
XX to viral infection. The AC1/CI genes are especially from BGW,  
XX tomato mottle virus or tomato yellow leaf curl virus (see T93292-93)  
XX and encode polypeptides (see W34334-35) that have mutations in the  
XX highly conserved DNA-nicking and/or the NTP-binding domains.

XX Sequence 353 AA;

Query Match 59.8%; Score 216; DB 18; Length 353;  
Best Local Similarity 58.6%; Pred. No. 3e-19;  
Matches 41; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

Qy 1 TLVWGEAAVGRSGAGCGCOTSDAAAEALNASSKEEALQIIREIKPKYLFQPHLNSNL 60  
Db 1 TLVWGEAAVGRSGAGCGCOTSDAAAEALNASSKEEALQIIREIKPKYLFQPHLNSNL 169

—

EFFC

## RESULT 8



```

R88872
ID R88872 standard; Protein: 359 AA.
AC R88872;
PA R88872;
DE 07-NOV-1996 (first entry)
XX Sardinian tomato yellow leaf curl virus mutated Rep protein (K227R).
XX Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
XX modification; mutation; viral replication; deficient; inhibition;
XX Sardinian tomato yellow leaf curl virus;
XX Sardinian isolate; STYICV; transgenic plant; P-loop; C1 protein;
XX AL1 protein.
XX OS Sardinian tomato yellow leaf curl virus.
XX
XX Key Location/Qualifiers.
XX Misc-difference 227
XX /note- "wild-type Lys has been replaced by Arg"
XX W09608573-A1.
XX
XX PD 21-MAR-1996.
XX PF 15-SEP-1995; 95MO-FR01192.
XX PR 15-SEP-1994; 94FR-0011040.
XX PA (CNRS ) CENT NAT RECH SCI.
XX Gronenborn B;
XX
XX WPI: 1996-179947/18.
XX N-PSDB; T12906.
XX Prodn. of virus-resistant transgenic plants - using mutated genomic
XX sequence from phytopathogenic DNA virus
XX
XX Disclosure: Fig 13; 93pp; French.
XX
XX Mutation of consensus amino acids in the NTP-binding site of
XX geminivirus Rep protein is used to produce replication deficient
XX transgenic plants that are resistant to, or tolerant of, the native
XX virus. The present sequence is a mutant form of the Rep (or C1)
XX protein from the Sardinian isolate of tomato yellow leaf curl virus
XX (STYICV) in which the wild-type Lys227 residue has been changed to
XX an Arg residue; transgenic Nicotiana benthamiana plants generated by
XX transformation with the mutated virus were not resistant to STYICV.
XX In contrast, plants transformed with a virus in which Lys227 had been
XX replaced by Ala were found to be resistant.
XX
XX Sequence 359 AA:

Query Match 59.8%; Score 216; DB 17; Length 359;
Best Local Similarity 56.5%; Pred. No. 3.1e-19;
Matches 39; Conservative 11; Mismatches 19; Indels 0; Gaps 0;
Oy 2 LWNKEAAVDCRSARGCCTNDAAAEALNASSKEEALQITREKIPKYLQFHLNLSL 61
Db 111 lwdgtfdgdsrgsgqdaayakagksqaldvkelaprdyvlhfnlnslid 170
Oy 62 RIFDKTPEP 70
Db 171 kvfivqpap 179
XX
XX RESULT 9
XX P70407
XX ID P70407 standard; Protein: 353 AA.
XX

```

```

AC P70407;
XX 02-MAY-1991 (first entry)
XX ORF 4 gene product of Bean Golden Mosaic virus.
XX Plant vector.
XX
XX OS Bean golden mosaic virus.
XX
XX JF61257186-A.
XX 14-NOV-1986.
XX
XX PF 10-MAY-1985; 85JP-0098108.
XX PR 10-MAY-1985; 85JP-0098108.
XX
XX (TEIJ ) TEIJIN KK.
XX WPI: 1987-159652/23.
XX N-PSDB; N70630.
XX
XX New DNA and hybrid DNA - used for recombinant vector of plants.
XX Disclosure: Fig 6; 24pp; Japanese.
XX
XX The sequence encoding this protein may be taken along with the -a
XX DNA sequence that encodes a stable resistance gene, and used to create a
XX recombinant plant vector.
XX See also N70629.
XX
XX Sequence 353 AA:

Query Match 59.3%; Score 214; DB 8; Length 353;
Best Local Similarity 58.8%; Pred. No. 3.1e-19;
Matches 41; Conservative 13; Mismatches 16; Indels 0; Gaps 0;
Oy 1 TLWNGEAAVDCRSARGCCTNDAAAEALNASSKEEALQITREKIPKYLQFHLNLSL 60
Db 110 tlewgfdgfdgsrsgsgqsandsyakaInadsesaitlkeqpkdyvlqhnlrsl 169
Oy 61 DRIFDKTPEP 70
Db 170 eriffkvpep 179
XX
XX RESULT 10
XX W34336
XX ID W34336 standard; Protein: 361 AA.
XX AC W34336;
XX
XX 27-APR-1998 (first entry)
XX
XX Tomato mottle virus AC1 protein.
XX
XX Geminivirus; TOMOV; AC1 gene; transdominant mutation;
XX transgenic plant; disease resistance.
XX
XX Tomato mottle virus isolate Florida.
XX W09739110-A1.
XX
XX 23-OCT-1997.
XX 15-APR-1997; 97WO-US06300.
XX 16-APR-1996; 96US-0015517.
XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.
XX (WISC ) WISCONSIN ALUMNI RES FOUND.
XX

```

XX PI Ahlquist PG, Hanson SF, Lou HT, Maxwell DP, Stout JT;  
 XX WPI: 1997-526447/48.  
 XX N-PSDB: T93284.  
 XX Transgenic plants expressing geminivirus AC1 and CI wild-type and  
 PT mutant genes - have increased resistance to geminivirus infection  
 PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean  
 PT golden mosaic geminivirus  
 XX Example 3.4: Page 57-58; 132pp; English.  
 XX This protein comprises the wild-type AC1 protein of tomato mottle  
 CC virus (ToMoV), a geminivirus that has a bipartite genome. The AC1  
 CC gene (see T93294) must be expressed for efficient replication of  
 CC the two genomic components, DNA-A and DNA-B. The AC1 protein has a  
 CC DNA binding site specific to the DNA-A common region, a DNA nicking  
 CC activity, and an NTP binding activity. The invention involves  
 CC production of transgenic plants containing DNA comprising AC1 or CI  
 CC genes with geminiviral replication during infection. Such transgenic  
 CC plants are resistant to viral infection. The AC1/CI genes are  
 CC especially from ToMoV, tomato yellow leaf curl virus or bean golden  
 CC mosaic geminivirus (see T93282-93) and encode polypeptides (see  
 CC W34324-35) that have mutations in the highly conserved DNA-nicking  
 CC domain and/or the NTP-binding domain.  
 XX Sequence 361 AA;

Query Match 59.3%; Score 214; DB 18; Length 361;  
 Best Local Similarity 55.7%; Pred. No. 5.5e-19;  
 Matches 39; Conservative 16; Mismatches 15; Indels 0; Gaps 0;  
 QY 1 TLVGEAAVDGSRGGCQTSNDAAAEALNASSKEALOITREKIPKYLFOFNLNSL 60  
 Db 110 tiewdfigdrgsrggqsandysakalnasvqslavireeqpkfivlqnhirsnl 169  
 QY 61 DRIFDKTPEP 70  
 Db 170 erifakepep 179

RESULT 11  
 W34324  
 ID W34324 standard; Protein: 361 AA.  
 AC W34324;  
 XX W34324;  
 DT 27-APR-1998 (first entry)  
 DE Tomato mottle virus AC1 mutant ToMoV-AC1dml.  
 XX Geminivirus; ToMoV-AC1dml; AC1 gene; transdominant mutation;  
 KW transgenic plant; disease resistance.  
 XX Tomato mottle virus isolate Florida.  
 OS Synthetic.  
 XX W09739110-A1.  
 PN W09739110-A1.  
 PD 23-OCT-1997.  
 XX 15-APR-1997; 97MO-US06300.  
 XX 16-APR-1997; 96US-0015517.  
 PR (SEMI-) SEMINIS VEGETABLE SEEDS INC.  
 PA (WISC) WISCONSIN ALUMNI RES FOUND.  
 PI Ahlquist PG, Hanson SF, Lou HT, Maxwell DP, Stout JT;  
 XX WPI: 1997-526447/48.  
 XX N-PSDB: T93283.

DR WPI: 1997-526447/48.  
 DR N-PSDB: T93282.  
 XX Transgenic plants expressing geminivirus AC1 and CI wild-type and  
 PT mutant genes - have increased resistance to geminivirus infection  
 PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean  
 PT golden mosaic geminivirus  
 XX Example 3.4: Page 60-62; 132pp; English.  
 XX This protein comprises a transdominant lethal mutant, designated  
 CC ToMoV-AC1dml, of tomato mottle virus (ToMoV) AC1 protein (see  
 CC W34324). It is encoded by a mutant AC1 gene (see T93282) of  
 CC ToMoV. The AC1 gene is located in the DNA-A component of its  
 CC NTP-binding domains. The AC1 gene, see also T93294, must be  
 CC expressed for efficient replication of the two genomic components,  
 CC DNA-A and DNA-B, of the bipartite ToMoV genome. The invention  
 CC involves production of transgenic plants containing DNA comprising  
 CC geminivirus AC1 or CI wild-type or mutant sequences that negatively  
 CC interfere in trans with geminiviral replication during infection.  
 CC Such transgenic plants are resistant to viral infection. The  
 CC AC1/CI genes are especially from ToMoV, tomato yellow leaf curl  
 CC virus or bean golden mosaic geminivirus (see T93282-93) and encode  
 CC polypeptides (see W34324-35) that have mutations in the highly  
 CC conserved DNA-nicking and/or NTP-binding domains.  
 XX Sequence 361 AA;

Query Match 59.3%; Score 214; DB 18; Length 361;  
 Best Local Similarity 55.7%; Pred. No. 5.5e-19;  
 Matches 39; Conservative 16; Mismatches 15; Indels 0; Gaps 0;  
 QY 1 TLVGEAAVDGSRGGCQTSNDAAAEALNASSKEALOITREKIPKYLFOFNLNSL 60  
 Db 110 tiewdfigdrgsrggqsandysakalnasvqslavireeqpkfivlqnhirsnl 169  
 QY 61 DRIFDKTPEP 70  
 Db 170 erifakepep 179

RESULT 12  
 W34325  
 ID W34325 standard; Protein: 361 AA.  
 AC W34325;  
 XX W34325;  
 DT 27-APR-1998 (first entry)  
 DE Tomato mottle virus AC1 mutant ToMoV-AC1dml.  
 XX Geminivirus; ToMoV-AC1dml; AC1 gene; transdominant mutation;  
 KW transgenic plant; disease resistance.  
 XX Tomato mottle virus isolate Florida.  
 OS Synthetic.  
 XX W09739110-A1.  
 PN W09739110-A1.  
 PD 23-OCT-1997.  
 XX 15-APR-1997; 97MO-US06300.  
 XX 16-APR-1997; 96US-0015517.  
 PR (SEMI-) SEMINIS VEGETABLE SEEDS INC.  
 PA (WISC) WISCONSIN ALUMNI RES FOUND.  
 PI Ahlquist PG, Hanson SF, Lou HT, Maxwell DP, Stout JT;  
 XX WPI: 1997-526447/48.  
 XX N-PSDB: T93283.

Query Match	59.3%	Score 214	DB 18	Length 361
Best Local Similarity	55.7%	Pred. No. 5	5e-19	
Matches 39	Conservative 16	Mismatches 15	Indels 0	Gaps
QY	1	TLVGEAAVDGRSGCGQTSNDRAAPALNASSKEALQIQREKIDPEKYLQFNLSNL 60		
Dy	110	lhwgdqfkgrrsgygggaqdsykalnasavgsalavfrrdeqkdvqimlrsl 169		
Qb	61	DRTEFKTPEP 70		
Db	170	erifakepep 179		
RESULT 14				
W34329				
ID	W34329	standard; Protein; 357 AA.		
AC	W34329			
XX				
DT	27-APR-1998	(first entry)		
XX				
XX		Tomato yellow leaf curl virus mutant CI protein PTKY104R #1.		
XX		Geminivirus: PTKY104R #1; CI protein; transdominant mutation;		
KW		transgenic plant; disease resistance.		
XX				
OS		Tomato yellow leaf curl virus strain Israel.		
OS		Synthetic.		
XX		W09739110-A1.		
XX				
PD		23-OCT-1997.		
XX				
PF		15-APR-1997; 97WO-US06300.		
XX				
PR		16-APR-1996; 96US-0015517.		
PA		(SEMI-) SEMINIS VEGETABLE SEEDS INC.		
PA		(WISC) WISCONSIN ALUMNI RES FOUND.		
XX				
PI		Ahlquist RE, Hanson SF, Luu HT, Maxwell DP, Stout JT;		
XX				
DR		WP1; 1997-526447/48.		
DR		N-PSDB: T93287.		
XX				
PT		Transgenic plants expressing geminivirus AC1 and CI wild-type and		
PT		CI mutant virus constructs to enhance resistance to leaf curl		
PT		P.g. tomato mottle virus, tomato yellow leaf curl virus or bean		

PT golden mosaic geminivirus

XX Example 4.5; Page 84-86; 132pp: English.

CC This protein comprises a transdominant lethal mutant of the tomato  
 CC yellow leaf curl virus (TYLCV) C1 protein (see W34337) that is  
 CC required for replication. It is encoded by construct pTYK225A #1  
 CC (see T93288) of an infectious clone of TYLCV and carries a single  
 CC mutation in the DNA-nicking domain. The invention involves  
 CC production of transgenic plants containing DNA comprising  
 CC geminivirus C1 or AC1 wild-type or mutant sequences that negatively  
 CC interfere in trans with geminiviral replication during infection.  
 CC Such transgenic plants are resistant to viral infection. The  
 CC AC1/C1 genes are especially from TYLCV, tomato mottle virus or  
 CC bean golden mosaic geminivirus (see T93282-93) and encode  
 CC polypeptides (see W34324-35) that have mutations in the highly  
 CC conserved DNA-nicking and/or NTP-binding domains.  
 XX Sequence 357 AA:

Query Match 58.7%; Score 212; DB 18; Length 357;  
 Best Local Similarity 65.6%; Pred. No. 9,7e-19;  
 Matches 40; Conservative 11; Mismatches 10; Indels 0; Gaps 0;

OY 4 MGEAVDGRSGGCGTNDAAEALNASSKEALQIIRKPEKYLQFHNLSNLDRI 63  
 Db 111 fgvsgldgrsargggqgsandayaealnssksealnllkekpkdyliqfhnlssnldri 170  
 OY 64 F 64  
 Db 171 f 171

RESULT 15

W34330 4 tomato yellow leaf curl virus mutant C1 protein pTYK225A #4.

ID W34330 standard; Protein; 357 AA.

AC W34330;

DT 27-APR-1998 (first entry)

XX Tomato yellow leaf curl virus mutant C1 protein pTYK225A #4.

XX Geminivirus: pTYK225A #4; C1 protein; transdominant mutation;

KW transgenic plant; disease resistance.

XX Tomato yellow leaf curl virus strain Israel.

OS Synthetic.

XX W09739110-A1.

XX W09739110-A1.

XX 23-OCT-1997.

XX 15-APR-1997; 97W0-US06300.

XX 16-APR-1996; 96US-0015517.

XX (SEMI-) SEMINIS VEGETABLE SEEDS INC.

PA (WISC) WISCONSIN ALUMNI RES FOUND.

XX Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;

DR WPI; 1997-526447/48.

XX N-PSDB; T93288.

XX Transgenic plants expressing geminivirus AC1 and C1 wild-type and

PT mutant genes... have increased resistance to geminivirus infection

PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean

XX golden mosaic geminivirus

XX Example 4.5; Page 88-90; 132pp: English.

CC This protein comprises a transdominant lethal mutant of the tomato  
 CC yellow leaf curl virus (TYLCV) C1 protein (see W34337) that is  
 CC required for replication. It is encoded by construct pTYK225A #4  
 CC (see T93288) of an infectious clone of TYLCV and carries a single  
 CC mutation in the DNA-nicking domain. The invention involves  
 CC production of transgenic plants containing DNA comprising  
 CC geminivirus C1 or AC1 wild-type or mutant sequences that negatively  
 CC interfere in trans with geminiviral replication during infection.  
 CC Such transgenic plants are resistant to viral infection. The  
 CC AC1/C1 genes are especially from TYLCV, tomato mottle virus or  
 CC bean golden mosaic geminivirus (see T93282-93) and encode  
 CC polypeptides (see W34324-35) that have mutations in the highly  
 CC conserved DNA-nicking and/or NTP-binding domains.  
 XX Sequence 357 AA:

Query Match 58.7%; Score 212; DB 18; Length 357;  
 Best Local Similarity 65.6%; Pred. No. 9,7e-19;  
 Matches 40; Conservative 11; Mismatches 10; Indels 0; Gaps 0;

OY 4 MGEAVDGRSGGCGTNDAAEALNASSKEALQIIRKPEKYLQFHNLSNLDRI 63  
 Db 111 fgvsgldgrsargggqgsandayaealnssksealnllkekpkdyliqfhnlssnldri 170  
 OY 64 F 64  
 Db 171 f 171

Search completed: February 3, 2001, 02:15:29  
 Job time: 3436 sec